DOCUMENT RESUME

ED 071 068 24 CS 000 351

AUTHOR	Coomber, James Elwood
TITLE	A Psycholinguistic Analysis of Oral Reading Errors Made by Good, Average, and Poor Readers.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	WRDCCL-TR-232
BUREAU NO	BR-5-0216
PUB DATE	Sep 72
CONTRACT	OEC-5-10-154
NOTE	122p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Context Clues; Grade 3; Psycholinguistics; Reading Ability; *Reading Skills; Sentence Structure; Syntax
IDENTIFIERS	*Error Analysis; Phoneme Grapheme Correspondence

ABSTRACT

Thirty third graders, divided into three equal groups, were used to determine the extent to which good, average, and poor readers depend upon two types of reading cues--Graphic features of word and of context. To hold error quantity differences constant, materials were chosen at different levels of vocabulary and syntactic difficulty. Each subject began reading first-grade level materials and continued through increasingly difficult readings until he made over twenty errors for every one hundred words he read. Errors were scrutinized for graphic-phonetic and syntactic resemblance to printed word, subject's dependence on grapheme-phoneme relationships, and location of error in response. Awareness of syntax was measured by an evaluation of the appropriateness of substitution errors to sentences in which they occurred and by an analysis of the position of errors in phrases. Conformity of multiword repetitions to phrase structure grammar was also noted. Results revealed that on longer words the good readers were more likely to err with a word that matched the stimulus in both initial and final positions. Syntactic measures revealed a high sensitivity to sentence structure. It was concluded that the readers were distinguished by differences in processing graphemes. (This document previously announced as ED 067 613.) (Author/HS)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Technical Report No. 232

A PSYCHOLINGUISTIC ANALYSIS OF ORAL READING ERRORS MADE BY GOOD, AVERAGE, AND POOR READERS

Report from the Project on Development of Instructional Programs: Wisconsin Design for Reading Skill Development

by James Elwood Coomber

Wayne Otto
Principal Investigator

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

September 1972

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Wisconsin R and D Center
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in The University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

STATEMENT OF FOCUS

Individually Guided Education (IGE) is a new comprehensive system of elementary education. The following components of the IGE system are in varying stages of development and implementation: a new organization for instruction and related administrative arrangements; a model of instructional programing for the individual student; and curriculum components in prereading, reading, mathematics, motivation, and environmental education. The development of other curriculum components, of a system for managing instruction by computer, and of instructional strategies is needed to complete the system. Continuing programmatic research is required to provide a sound knowledge base for the components under development and for improved second generation components. Finally, systematic implementation is essential so that the products will function properly in the IGE schools.

The Center plans and carries out the research, development, and implementation components of its IGE program in this sequence: (1) identify the needs and delimit the component problem area; (2) assess the possible constraints--financial resources and availability of staff; (3) formulate general plans and specific procedures for solving the problems; (4) secure and allocate human and material resources to carry out the plans; (5) provide for effective communication among personnel and efficient management of activities and resources; and (6) evaluate the effectiveness of each activity and its contribution to the total program and correct any difficulties through feedback mechanisms and appropriate management techniques.

A self-renewing system of elementary education is projected in each participating elementary school, i.e., one which is less dependent on external sources for direction and is more responsive to the needs of the children attending each particular school. In the IGE schools, Center-developed and other curriculum products compatible with the Center's instructional programing model will lead to higher morale and job satisfaction among educational personnel. Each developmental product makes its unique contribution to IGE as it is implemented in the schools. The various research components add to the knowledge of Center practitioners, developers, and theorists.

ACKNOWLEDGEMENTS

I am especially grateful to my advisor, Dr. Wayne Otto, and to the other members of the dissertation committee, Dr. Richard L. Venezky and Dr. Dale D. Johnson. Their advice and encouragement were invaluable in all phrases of the study. Despite their busy schedules, they were so willing to take time to discuss ideas and problems.

The willingness of Dr. Richard J. Smith and Dr. Peter A. Schreiber to serve on the oral defense committee is also appreciated. I also wish to thank Dr. Smith for his critique of the study in its early stages.

Other faculty who gave valuable advice include Dr. Robin Chapman, Dr. Carl Personke, Dr. Kenneth L. Dulin, and Dr. Lester Golub, now of Pennsylvania State University. Suggestions from Dr. Rose-Marie Weber of McGill University and from Dr. Andrew Biemiller of the University of Toronto also provided valuable guidance.

For assistance in statistical analysis and research design I wish to thank two of my colleagues at Concordia College, Dr. Gary Barko and Dr. James Ulness. Computer assistance from Mr. Ed Haertel at Wisconsin and from Mr. James Zum Brunnen of North Dakota State University was also greatly appreciated.

I am very grateful to the Wisconsin Research and Development Center for Cognitive Learning for financial support and for my use of Center facilities. Another real financial help during the time I was doing this research was the scholarship awarded me by the Board of College Education of the American Lutheran Church.

Appreciation is also extended to Mr. Robert Soderbloom and teachers in the Stoughton, Wisconsin, school system for their kind cooperation in the study.

Finally, I am grateful to my wife, not only for support and encouragement, but also for the many long hours she spent doing clerical work for the "cause." And I am also grateful to other family, friends, and colleagues who have expressed interest in my work.

Jim Coomber

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	viii
ABSTRACT	ix
CHAPTER	
I. INTRODUCTION	1
General Purpose of the Study	1
Review of the Literature	4
Early Studies of Oral Reading Errors	6
Recent Error Analyses	9
Rationale for the Hypotheses Tested	11
Contextual Cues	11
The Relationship of Graphic Cues and Contextual Cues	15
Use of Grapheme-Phoneme Correspondences	19
Measures of Use of Reading Cues	24
Determining Use of Grapheme-Phoneme Correspondences	24
Determing Use of Syntactic Context	26
Summary of Hypotheses	36
II. METHOD	41
Development of the Instrument	42
The Pilot Study	43
The Final Instrument	44
The Study	47
Subjects	47
Procedure	49
Data Gathered	51
Analysis	53
Measures on the Word Level	54
Measures on the Syntactic Level	59
Summary	64
III. RESULTS AND DISCUSSION	65
Results	65
Hypothesis 1	65
Hypotheses 2 and 3	79
Hypothesis 4	81
Hypothesis 5	84
Hypothesis 6	88
Summary	90

IV. SUMMARY, LIMITATIONS, AND IMPLICATIONS 91
Summary 91
Problem 91
Subjects 92
Procedures 92
Results 94
Limitations of the Study 97
Implications for Teaching 99
Implications for Future Research 103
Word Level 103
Syntactic Level 106

BIBLIOGRAPHY 109

LIST OF TABLES
IN THE TEXT

Table		Page
1	Mean Error Rate in Two Sets of Passages--All Subjects	45
2	Mean Error Rate in Two Sets of Passages--Good Readers Only	45
3	Agreement on Placement of Subjects by Teachers and Test Results	48
4	Performance Scores of Subjects on the Gates-MacGinitie Reading Achievement Test	49
5	Graphic Similarity Index Formula	57
6	Percentages of Long- and Short-Word Errors	67
7	Percentages of Error Types--Long Words	68
8	Percentages of Error Types 2' and 4' to Total Long Word Errors	70
9	Percentages of Error Types--Short Words	72
10	Percentage of Function Word Errors and Content Word Errors--Short Words	73
11	Percentages of Error Types on Short Word Errors--Function Words-Content Words	74
12	Percentages of Error Types on Short Function Words	75
13	Percentages of Error Types on Short Content Words.	75
14	Mean Graphic Similarity Scores	79
15	Overall Analysis of Variance Based on Mean Graphic Similarity Scores	80
16	Proportions of Grammatically Inappropriate and Non-Word Errors	82

Table

17 Overall Analysis of Variance Based on Proportions of Grammatically Inappropriate and Non-Word Errors . . .82

18 Location of Errors in Phrases 85

19 Percentage of Errors by Third of Clause 87

20 Conformity of Multi-Word Repetitions to Phrase-Structure Grammar 89

LIST OF FIGURES

Figure

1 Pike's Representation of Constituent Analysis 28

2 Tree Diagrams as Used in Transformational Grammar . . . 30

ABSTRACT

The oral reading errors of selected third-grade children were analyzed as a means of studying differences in reading strategy between good, average, and poor readers. The graphic-phonetic and syntactic appropriateness of errors to their printed stimulus words was considered. Reading materials were controlled for vocabulary and syntactic difficulty in order to hold the error quantity factor constant and also to consider syntactic competency without the effect of vocabulary difficulty. Thus readers of all levels of reading proficiency could be compared at similar levels of error rate as well as on the same passages.

The data revealed that the differences between readers of different levels of reading proficiency were on the graphic-phonetic level. The better the reader, the greater was the graphic-phonetic appropriateness of his oral reading errors to the printed stimuli in long word errors. This observation did not hold true with errors on short words, largely because of the word frequency factor. On the syntactic level, a study of the grammatical appropriateness of errors and the results of several psycholinguistic test of the data indicated no significant differences between the groups in their syntactic competence in reading.

Chapter I

INTRODUCTION

General Purpose of the Study

The purpose of this study was to find differences, if any exist, in the reading strategies of good, average, and poor readers[1] through an analysis of oral reading errors. The investigator examined the extent to which readers of each level of reading proficiency depended upon the two main types of reading cues--the graphic features of the word itself and the context in which the word appears--and the degree of success different readers experienced in using those cues. Substitutions were the center of focus because they provide the greatest amount of insight into reading strategy. On the word level, such mismatches make possible an examination of what part or parts of a word are the strongest cues for readers. On the sentence level, substitutions indicate reader awareness of grammatical constraints and phrase structure. Also, the fact that substitutions generally occur in greater frequency than any of the other oral reading errors made them a feasible subject

[1] For purposes of this study the term poor reader does not include the so-called non-reader, the child who is unable even to be tested for oral reading errors. The poor readers in the experiment possess at least minimal reading competence and can read orally materials written for children two years younger. It would seem that, in other oral reading studies in which reader ability has been a variable, the totally incompetent reader has not been included.

for analysis. Thus the study of patterns of oral reading errors can provide insight into reading strategy and reading deficiency.

Generally speaking, there are two explanations, not necessarily mutually exclusive, for differences in performance between readers of different levels of proficiency. First, one might regard the difference as merely a quantitative one. According to this view, espoused in Malmquist's study (1958), disabled readers make the same kinds of errors as good readers, but in greater number. An alternate position claims a qualitative, as well as a quantitative, difference in the errors of different types of readers. If this is the case, as some recent work by Weber (1970b) seems to indicate, one would expect to find a weakness, or perhaps a syndrome of weaknesses, characteristic especially of the reading performance of poor readers. It is also possible that no such syndrome exists, that each reader, regardless of his reading proficiency, has his own unique pattern that cannot be generalized to a group. The writer dealt with this question, using techniques of linguistic analysis that have come into use since Malmquist's inquiry.

Few studies of oral reading errors have been devoted to a comparison of the errors of good, average, and poor readers beyond the first-grade level. Malmquist's study (1958) was an attempt to reveal specific reading behaviors characteristic of good, average, and poor young readers in Sweden. All subjects read the same passages, and the experimenter recorded and analyzed their errors. Malmquist concluded that there are not weaknesses in reading that are characteristic of only the poor readers. Rather, all readers tend to make the same patterns of errors, but in varying quantities. Malmquist, however, claimed to

do little more than tabulate error types--the number of substitutions, the number of omissions, and the like. There was no attempt to evaluate degrees of appropriateness of errors--semantic, syntactic, graphic, and phonemic--in the sense that K. S. Goodman (1969), for example, has proposed. These techniques have been used in none of the large-group error analyses beyond the first grade. Weber's analysis of grade one reading errors (1970b) provides at least a suggestion that, by using some of the measures developed since the Malmquist study, it might well be possible to find qualitative, as well as quantitative, differences between the errors of more mature readers of diverse reading ability. Hence this study was addressed to the problem posed in the Malmquist inquiry, but more refined evaluative techniques were employed.

Thus it was believed that comparing readers at different levels of reading proficiency with error quantity held constant might reveal qualitative differences. Examining errors of all readers at specific levels of proficiency--for example, at the point at which all subjects miss five words out of a hundred--should have the effect of holding error quantity constant. At these levels of proficiency, then, it should be feasible to examine the data for qualitative differences between the readers. If the difference between the reading errors of good, average, and poor readers is merely a matter of quantity, one would expect proportionately about the same number of errors at any given level of proficiency and no significant difference in degree of approximation to the printed word and in syntactic and semantic appropriateness to the context. If, on the other hand, there are qualitative

differences, some diversity in types of errors or different degrees of appropriateness to context would be expected. Also, such controls for error quantity can allow for a study of shifts in error patterns as readers advance from easy to more difficult material. The extent to which proportions of error types--substitutions, omissions, and additions--change with reading difficulty can also be studied, as well as differences in this regard between readers of different levels of reading proficiency.

In addition to finding distinctions between good and poor readers, the study revealed characteristics of the reading process in children at the third-grade level. Specifically, the writer noted the extent to which readers rely on grapheme-phoneme correspondences and on contextual cues in reading selections of different levels of difficulty.

Review of the Literature

Through studies of oral reading errors, one can gain valuable insights into the reading process and into the strategies used by individual readers. It is, of course, difficult to trace the perceptual and mental processes that occur as an individual reads. The only means of observing such inner workings is through a study of their external manifestations. A mistake reveals a flaw in the process, a failure of the response to wholly match the stimulus. By studying these instances of mismatch between stimulus and response, it is possible to gain insights into the processes used in reading; one can note both where stimulus and response match and where they fail to match. Through such studies one can obtain diagnostic information on individual children as well as trace the development of reading skills in groups of readers.

With this information it is possible to determine the degree to which readers at different ages and at different levels of proficiency depend on the different types of cues. Through studies of reading errors, researchers have demonstrated that readers do use certain cues, for example, letters and letter combinations, and semantic and syntactic context, quite consistently (K. S. Goodman, 1967; Y. Goodman, 1967). From such analyses of reading behavior it is possible to explain why children make the reading errors they do. The fact that specific strategies are so apparent and so reasonable in reading errors prompted K. S. Goodman to label them "miscues," in contrast to the notion that such mistakes result merely from ignorance or carelessness. Page (1970) especially objects to the practice of calling mismatches "errors", pointing out the subjective nature of such a label:

> To attempt to draw the value conclusion that a miscue is an error from the descriptive information that a miscue was generated without intervening value warrants or reasons why the miscue should not have been generated poses a complex but obvious logical fallacy. (p. 20)

Elsewhere Page even suggests that poor readers are labelled as such because of an "unwarranted value judgment." Few reading specialists would share Page's benign attitude toward reading errors and his emphasis on the subjectivity of teacher judgments of reading performance. Nevertheless, his view at least encourages discrimination between types and degrees of mismatch and in this way seems a helpful qualification of Gilmore's (1947) branding of such errors as "undesirable characteristics of oral reading." (p. 152)

Oral reading studies are based on the assumption that oral reading ability is a fairly close reflection of silent reading ability and that

performance in the observable oral medium can be generalized to the largely unobservable silent medium. Weber, in her review of the literature on oral reading studies (1968), considers the correlation between the two modes of reading in the research reported by Swanson and by Gilmore. Swanson (1937) demonstrated the existence of a positive relationship between oral reading ability and silent reading ability as measured by comprehension questions and identification of words presented by a tachistoscope, the latter task being considered the equivalent of silent reading. Gilmore (1947) obtained moderately high correlations between oral reading performances and scores of reading comprehension, both silent and oral, for subjects in grades one through eight. Although both studies show a high correlation between oral performance and these overt manifestations of silent reading, Weber concludes that there is no truly relevant research based on "common elements in the reading task itself." She makes no suggestions as to possible methods for obtaining such direct information.

Weber has made a comprehensive review (1968) of the literature on oral reading. The reader is referred to her compilation for summaries of studies on correlations between oral reading performance and age, sex, intelligence, and other pertinent factors. The Gilmore dissertation (1947), though dated, is a particularly valuable source of information on the correlations between oral reading performance and other relevant factors. Only the inquiries that are directly relevant to the present purpose are considered in detail in this chapter.

Early Studies of Oral Reading Errors

Oral reading studies have proven to be a valuable means of studying

reading behavior, particularly in the last few years. The approach is not a new one; it dates as far back as 1930. The early studies involved mainly counting and classifying errors for the purpose of evaluating reading difficulties (Payne, 1930; Madden and Pratt, 1941). Some of these researchers were concerned with quality of oral communication in general, observing such factors as expression, word-by-word reading, as well as ability to decode (Duffy, 1935; Daw, 1938). A considerable amount of this research, probably triggered by the Orton hypothesis,[2] focused on frequency of reversals. While the techniques used seem gross by comparison to the approaches used today, these early researchers made substantial contributions to our knowledge of the reading process. For instance, it was Payne (1930) who first noted what Goodman has stressed more recently--that children's oral reading errors are not without patterns. Payne found that frequency of occurrence of the substituted word and similarity of configuration and sound seems to govern the choice of substitued words. In trying to account further for a child's choice of substitutions, she reported, "When confronted with an unknown word, a child has a tendency to call it the word in his sight vocabulary which is most like it in sound and appearance." (p. 146)

Bennett (1942), in her remedial program with poor readers in the middle grades, carried out the first study in which similarity of stim-

[2] Samuel Orton, a neurologist, speculated in the 1930's that many cases of reading disability are due to a lack of consistent dominance of one hemisphere of the brain over the other. Lack of clear dominance, Orton suggested, is manifested by a disproportionately large number of reversal errors in oral reading.

ulus word and erroneous response was evaluated in a systematic way. She devised a scale for comparing which positons of the printed word and the mistaken response were similar--initial, medial, or final--along with several subclassifications. Her data indicated that the beginning and ending of a word are its most salient features, with the first letter twice as likely to be correct in an erroneous response as the last letter. This finding has since been corroborated in more carefully controlled studies (Marchbanks and Levin, 1965; Williams, Blumberg and Williams, 1970; and Venezky, Chapman and Calfee, 1971). Weber (1970b), in her consideration of oral errors, noted a similar tendency. In addition to salient letters and letter groups, Bennett discovered that context, too, played a significant role, as the children tended to substitute syntactically and semantically appropriate words. Bennett also found, in support of Payne's earlier finding, that children tend to substitute a meaningful word from their lexicons for a difficult word. Bennett's subjects missed a total of two hundred words in an oral reading assignment. All subjects made the same substitution for all except eighty-two of the two hundred words missed.

In a more recent study, Clay (1968) also examined appropriateness of error to context in her study of the oral reading of New Zealand first-graders. She found that the bulk of all words substituted--seventy-nine per cent--belonged to the same form class. Similarly, Weber (1970b) discovered that nine-tenths of the substitutions made by her first-grade subjects were "grammatically appropriate" to the point in the sentence where the error occurred. According to the reports by

Clay and Weber, the difference between high and low ability groups is negligible as far as appropriateness of words to the context is concerned.

Recent Error Analyses

In recent years new techniques have been introduced in error analysis, especially in the work of Allen (1969), K. S. Goodman (1967, 1969), Y. Goodman (1967), and Weber (1970b). Hypothesizing that the reader draws upon specific cues in reading, K. S. Goodman (1969) proposed that errors be analyzed on a number of linguistic levels--graphic, phonetic, syntactic, and semantic--to more precisely pinpoint the cues to which the reader responds. Thus he devised a set of eight or nine-level rating scales, any or all of which can be applied to any reading error for judging graphic, phonetic, syntactic, and semantic similarities. Weber (1970b) refined evaluation of similarity on the word level with her Graphic Similarity Index. Each substituted word is compared with its stimulus by means of a formula, which takes into account the total number of letters and pairs of letters that stimulus and response have in common, the length of the two forms, and resemblance of first and last letters. Different combinations earn different numbers of points so that each substitution can be evaluated according to a numerical score.

Both Weber (1970b) and Biemiller (1970) have characterized the development of the reading process in first-graders through the measures described above, comparing children's dependence on graphic and contextual clues. Both found an increasing dependence on graphic cues as children progressed in reading through grade one. Through the Gra-

phic Similarity Index described above, Weber noted that the errors of better readers were higher in graphic similarity to the printed words and that this graphic similarity of errors increased throughout the school year. On the other hand, the mean number of contextually appropriate errors dropped somewhat during the year. According to Biemiller, after a first-grader reaches a certain level of proficiency, his errors demonstrate some flexibility in word analysis. As he reads relatively easy material, he generally uses context. As the material becomes more challenging, there is a greater reliance on graphic cues. The readers who exhibit this flexibility--the ability to use graphic contextual cues appropriately--turn out to be the better readers.

The type of progress noted by Weber and Biemiller seems to conflict with studies of young readers over a longer period. Ilg and Ames (1950), in a longitudinal study of reading as children progressed from age five to age nine, observed that most children move from a preponderance of substitutions with graphic similarity at age five or six to a tendency to substitute words of the same meaning at age nine. Schale (1964) also considers errors in oral reading according to age, but from grades two through twelve. She observed no change in proportionate number of substitution errors. This apparent discrepancy with Ilg and Ames may be due to a reversal after first grade in the tendency toward greater graphic similarity of stimulus and response that Weber and Biemiller noted. Schale's findings can be explained by the fact that she was counting only the number of substitutions and other errors, rather than assessing the degree of similarity between stimulus and response.

As was explained above, the level of difficulty of the reading material is a variable to be explored in the present study. The two researchers (Schummers, 1956; Schale, 1966) who have considered this variable at the third grade level did not discover any dramatic difference in proportion of error types as children read materials of varying levels of difficulty. Substitutions did not vary in proportion with other types of errors according to the difficulty level of the reading. The reading proficiency of the subjects was not a variable in either study. These relatively gross measures used yielded little more than a count of error types rather than a measure of the resemblance of error and printed word. Biemiller's investigation (1970) of first-grade oral reading errors provides a hint that finer measures might reveal a shift in reading strategy in more mature readers as the prose becomes more difficult.

Rationale for the Hypotheses Tested

Evidence from some of the studies reviewed above provides a basis for explaining the relationship between contextual and graphic cues in the reading act. The limited evidence available suggests that when more mature readers are perusing relatively easy material, they use primarily contextual information--both semantic and syntactic--in processing words. But an increase in the difficulty of the reading task tends to produce a greater dependence on phonic skills.

Contextual Cues

The speaker of the language makes use of what some linguists have called "sequential constraints." Since the language-user knows the structure of the language, he expects to find certain word orders but

not to find others. For example, in English, subject usually preceeds verb, direct object generally follows the verb, articles may preceed but not follow the nouns they modify, and adjectives may preceed nouns but may not preceed pronouns. To illustrate the importance of word order in English, one might rearrange units of a simple sentence and note the confusion that results. The old man caught the biggest fish. Rearranging subject and predicate results in a completely opposite meaning in the sentence: The biggest fish caught the old man. Moving the verb of the original sentence to the end of the sentence does not change the meaning, but the resulting structure does not sound normal: The old man the biggest fish caught. Transposing article and adjective makes comprehension difficult: Old the man caught biggest the fish. Likewise, one might transpose adjective and noun to produce a sequence more characteristic of word order in French: The man old.... Other changes in word order could be cited, but these illustrations should suffice to show that the user of language, including the reader, expects certain sequences in sentences. Use of context in reading is a matter of responding to language characteristics such as word order.

Few users of the language are consciously aware of these constraints, nor of the more complex relationships that exist between nouns and verbs in sentences. For example, the speaker intuitively knows that some verbs require objects:

Jason resembles his father.
*Jason resembles.

*unacceptable string

Some verbs may or may not take an object:

The workers demonstrated against their oppressors.
The saleslady demonstrated the new eggbeater.

Qualities such as animate or inanimate, human or non-human, when applied to verbs, exert further constraint on word relationships. Lees, for example, has devoted much of his Grammar of English Nominalizations (1965) to an accounting of these relationships. Among the verb characteristics that he points out are that some verbs require animate or human subject or object. Verbs like reward and admire, for instance, take human subjects:

Sarah admired the house.
*The tree admired the house.

Some verbs may have either animate or inanimate subjects but require animate objects:

The massacre convinced me that the war is immoral.
McGovern convinced me that the war is immoral.
*McGovern convinced the machine guns that the war is immoral.

Many other examples of such constraints could be cited. All of them are an accepted part of our language.

Another important quality of language noted by early linguists, such as Bloomfield (1933), is the redundancy of language. Several structures might provide syntactic and semantic cues for only one item of knowledge. For example, as many as three or four morphemes may indicate plural subject, as is the case with both bound and free morphemes in the following sentences:

Those girls are eating their dinner.
The robins are leaving their nest.

In the following sentence, two morphemes indicate progressive aspect:

Mailer is writing his fourth novel.

K. S. Goodman (1966), in applying the redundancy principle to reading behavior, points out that, because redundancy is so common in our language, the reader can often overlook a letter, pattern, or an entire word but nevertheless fully understand an utterance. It is simply a matter of "filling in" certain gaps.

In addition to these syntactic expectations there are semantic ones, defined largely by the reader's world of experience and his assimilation of those experiences. Note the following two sentences:

Anna let the cat out. Then she telephoned her boyfriend.

She, being a pronoun, could refer either to cat or to Anna, but any child who is familiar with the telephone and, perhaps, with the abilities and limitations of cats, knows that the actor in the second sentence is Anna and not the cat. The reader's knowledge in general certainly determines his perception of the relationship of various parts of the sentence to each other. But, when meaning cues are insufficient, unclear references may be the result, as in the following two sentences:

As the Texans raised their rifles, the Mexican troops advanced. Then they attacked the Alamo.

Inadequate cuing results in a threefold ambiguity in the word they. Either the Texans or the Mexicans could be attacking, or it may be that the Texans and Mexicans in consort are trying to route a common enemy. Many other semantic relationships relevant to the reading process could be cited.

While linguists have written volumes of observations and analyses in their efforts to account for the various types of constraints and redundancies in the language, the vast majority of speakers of English are not conscious of these factors. Yet they observe the constraints almost without exception. Most children master these "rules" at an early age, before they begin school. Brown and Bellugi (1964) found that by the age of three the children they observed used all of the main types of simple sentences and revealed a consciousness of constituent structure. Thus the child who is learning to read already has a considerable asset in his favor for learning to read: an awareness of the structure of English.

The Relationship of Graphic Cues and Contextual Cues

As the reader encounters more challenging materials, his dependence on graphic cues should increase; the syntactic and semantic cues simply do not yield enough information to "unlock" certain words. This is not an all-or-none matter. Most likely the reader is, at all times, responding to the four types of cuing--graphic, phonetic, syntactic and semantic. As the material becomes more challenging, there is probably a gradual shift from greater reliance on context cues to greater reliance on word attack skills. A greater number of words are not recognized on sight and must be analyzed carefully. As difficult words become more frequent, usable context will gradually break down, particularly if the reader encounters words that are not even in his vocabulary. Usable context might ultimately disintegrate to the point that the reader might as well be calling out strange words as in a word list--if anyone cared to carry the matter to such an extreme. At a

less extreme level, however, the presence of a few difficult or strange words will slow the reader to the point at which he must attend carefully to individual words and to the individual letter and letter combinations within words. In relatively easy oral reading tasks, proficient readers demonstrate these tendencies by omitting articles and some word endings, but their accuracy in responding to these structures should increase as the material becomes more difficult.

One study of subvocalization would seem to indicate that the relationship suggested above is valid. Hardyck and Petrinovich (1969) showed that subvocalization in high school and college subjects increased in proportion to the difficulty of the reading material. They found that readers who normally did not subvocalize, did so when they read difficult material. They compared subvocalization with shifting into low gear in a car; in both situations the performer is drawing upon an added source of power. While Hardyck and Petrinovich did not state or suggest it, it seems likely that increased subvocalization might correspond to a greater dependence on the graphic display. That is, subvocalization may be a sign that the reader is "sounding out" much of what he is reading.

One question of this study is the extent to which readers--good, average, and poor--switch strategies in this way. This writer suggests that all readers move from dependence predominantly on contextual cues to greater dependence on the word's visual components as the reading task demands it. This study is an attempt to discover what it is that the average reader does that the poor reader fails to do--and that good readers do so well. The problem may be a matter of proficiency in

switching strategies--from using primarily context to attending more closely to the parts of the words themselves. The good reader might be one who has achieved greater than average flexibility in using reading strategies while the poor reader is one who has failed to master this skill. The poor reader, provided he possesses at least a minimal level of reading competency, might well be nearly the equal of average or good readers as long as he can effectively use context--assuming he has had at least some experience with the subject matter under consideration. It seems likely that the good reader, on the other hand, has the same command of contextual cues for relatively easy reading but in addition demonstrates proficiency in processing graphemes when the occasion demands it.

Some support for distinguishing between good and poor readers on this basis comes from several recent studies of the errors of first-grade readers, cited above. In his study of oral reading errors, Biemiller identified three stages in the beginning reading process. He found that initially children rely almost wholly on semantic and syntactic context for attacking words. After some time, a period varying considerably in length among the youngsters, Biemiller observed a tendency toward "non-responses," when children refused to respond to difficult words. He regarded this as evidence of the fact that they were focusing on the words and their parts. Apparently they were aware of the increased difficulty of the task and their inability to cope with it. Most children, after a few weeks or more in the non-response stage, entered a third stage, in which they demonstrated flexibility in method of word attack. When material increased in difficulty--specifically to

the point that the child was making errors on five to ten out of every hundred words read—errors showed that the child tended to shift from use of contextual cues to use of graphic cues. Some children never progressed beyond the first stage during grade one; these turned out to be the poorer readers. It should be pointed out that Biemiller's method of determining dependence on graphic similarity was rather gross; if the first letter of stimulus and response were alike, he inferred the use of graphic cues. Nevertheless, his findings are highly suggestive.

Biemiller's findings on use of context are in agreement with those of Clay and Weber. Clay (1968) found that good readers were not significantly superior to poor readers in their ability to substitute words of the same form class. In fact, poorer readers surpassed superior readers in their substituting grammatically appropriate single words. Weber (1970b) found little difference between high and low groups in substituting words appropriate to the context.[3] The percentage of grammatically appropriate errors for both groups dropped through the year, but the low groups experienced the greatest drop—about ten percentage points. One might ascribe this observation to the fact that, as poor readers encountered increasingly more difficult material without having mastered the prerequisite phonic skills, usable context decreased for them. It would be necessary for them to give greater attention to the graphic display--likely at the expense of attending to grammatical con-

[3] Neither Clay nor Weber used statistical analysis on their data. Thus it is not possible to characterize their findings in terms of statistical significance.

straints. The degree to which good, average, and poor readers beyond the first-grade level differ in their tendency to substitute contextually appropriate words is not known. Bennett's finding (1942) that a vast majority of the errors of poor readers in the intermediate grades fit the context in which they occurred provides some basis for suggesting that differences between the reading ability groups in this regard might be minimal.

Thus it seems predictable that in an oral reading task at the third-grade level there will be little difference in appropriate use of context across reading ability groups--at given levels of reading proficiency for all three groups. While the studies cited offer some reasons for expecting this tendency, one must be careful in generalizing from the findings of first-grade studies--those by Biemiller, Clay, and Weber--to reading behavior in the present study. Whether the context-graphic cue relationship holds at more advanced levels, such as in grade three, remains unproven. Too, only the Biemiller investigation considered difficulty of reading material as a variable, and not in relation to the diverse abilities of the subjects. Further research might reveal the extent to which these factors are present at the other grade levels.

Use of Grapheme-Phoneme Relationships

The reader's use of grapheme-phoneme relationships is determined by the extent to which oral reading errors of good, average, and poor readers conform to the graphic display. The writer does not mean to imply that some children are strictly context-readers while others use

primarily graphic cues. In fact, it might well be that poor readers use graphic cues very considerably, perhaps even to the same extent that good readers do. But it seems that they are most successful in operating from context. Why are they unable to use effectively the grapheme-phoneme relationship? Two basic reasons come to mind. Perhaps they fail "to see the whole picture" and respond to just a part of the stimulus, presumably the most salient feature. That is, if there is a more predominant part of a stimulus, the poor reader would be expected to perceive it, if nothing else. Secondly, poor readers may be responding to irrelevant or misleading parts of the stimulus, such as word shape.

Several studies have demonstrated that the initial part of the word is its most salient feature. Bennett (1942), in her analysis of location of errors in words, found "almost a two to one chance" that the beginning of the word is a more potent cue than the ending. Weber (1970b) obtained similar results in her oral reading error study of first-graders. Marchbanks and Levin (1965), using three-letter and five-letter synthetic words in a study with kindergarten and grade one children, found that the initial letter was the most powerful cue in word identification. The final letter was the second most potent part of the word. These researchers suggested that their results might be explained by the theory that the first cue or the last cue of a stimulus is the strongest. Their alternative explanation is that the first and last letters of a word might be the more obvious letters because each is surrounded on the side by mere space. Williams, Blumberg, and Williams (1970) replicated the Marchbanks and Levin study in a "matching

to sample" task and obtained essentially the same results with kindergarteners and first graders.

Venezky, Chapman, and Calfee (1971) investigated position of error in children's pronunciation of synthetic words, with age as a variable. Subjects were children in the upper and lower quartiles in grades two, four, and six. Invariant consonants (e.g., b, d, l, m, and n) were controlled in three positions of the trigrams--initial, medial, and final. The difference in error rate for the initial position was not significant between groups at any of the grade levels, but there was a significant difference between the two proficiency groups in medial and final positions. This intergroup difference diminished with age. For both proficiency groups, the initial position proved to be the most salient cue. Contrary to Marchbanks and Levin (1965), Venezky, Chapman, and Calfee did not find significant differences in cue strength between medial and final positions in the trigrams in either proficiency group, though there was a tendency for more errors to occur in the final position. The apparent conflict between these two studies is likely due to the fact that in the earlier study the children were asked to make visual discriminations while in the latter study they pronounced the synthetic words. Age differences might also be a factor; Marchbanks and Levin's subjects were in kindergarten and grade one while the youngest subjects in the study by Venezky, Chapman and Calfee were in grade two. The differing results between the kindergarteners and the first graders may well reflect the tendency for

children to go beyond the most salient cues in word recognition as they mature or as they receive instruction in reading. Bennett (1942) found that as the school year progressed and her subjects received tutoring in reading, their oral reading errors showed a change in pattern, specifically in the location of errors in words. Errors in which only the initial parts of stimulus and response matched decreased considerably. On the other hand, she reported increases in the number of errors in which the middle positions of stimulus and response resembled each other.

These findings, then, seem to suggest that one difference between the three types of readers is a matter of choosing the most powerful cue--namely, the initial letters and perhaps to a lesser extent the final letters. In an oral reading error study at the third-grade level, this perceptual habit should be manifested by a tendency, especially among poor readers, to make a large number of erroneous responses that match the stimulus only in the initial position. The tendency for a match in the final position, according to the studies by Marchbanks and Levin (1965) and Williams, Blumberg, and Williams (1970), should be the next to strongest, though the results obtained by Venezky, Chapman and Calfee do not support this expectation.

Word length might affect oral reading error patterns. While better readers presumably make more graphically and phonetically correct errors on long words, this tendency might not be observed in errors on short words. It seems likely that the poor reader might analyze each word--even function words--carefully while the good readers are more inclined to respond to semantic and syntactic cues and thus make errors

that are not high in graphic and phonetic similarity. Poor readers might be able to decode short words more accurately than long words simply because the graphic display is less complicated. Thus it was expected that children might attack short words and long words differently. The nature of this difference could not be predicted at this point.

The possibility that dependence on the wrong kinds of cues keeps poor readers from dealing effectively with graphemes was suggested in studies done over three decades ago. In trying to account for the substitutions children make, Payne (1930) suggested sound and word shape as influences on many erroneous word choices. A child encountering a strange word, she reasoned, tends to respond with a word from his lexicon that has a similar configuration or sound. Swanson (1937), examining various facets of oral reading in students at the University of Iowa, reported that sixty per cent of all substitutions made by poor readers displayed similar word shape to their printed stimuli. On the contrary, only forty-two per cent of the good readers' responses resembled the stimuli in word shape. Generalizing the findings of studies using college-age subjects to the reading process in primary grade children seems risky indeed. But if university students resort to using such weak cues, it seems likely that they might have developed the habit at an early age--and never dropped it. It should be noted, however, that Marchbanks and Levin (1965) and Williams, Blumberg, and Williams (1970) found configuration the least used cue for word identification in their kindergarten and grade one subjects. In fact, the authors of the latter study rejected the notion that configuration is

a potent cue even among kindergarteners who do not know the alphabet.

Measures of Use of Reading Cues

Detailed discussion of the linguistic measures used in the study is postponed until the chapter on methods. However, the hypotheses stem from the measures themselves, so a general explanation of them is included before the hypotheses are presented. There are several methods used for assessing reader's use of the two major cue systems--grapheme-phoneme correspondences and context.

Determining Use of Grapheme-Phoneme Correspondences

To what degree do mismatches resemble the printed word? The writer attempted to identify the differences in error patterns between good, average, and poor readers at the third-grade level by examining the extent to which each type of reader uses contextual and graphic cues under similar and different conditions of reading difficulty. Two means of examing dependence on grapheme-phoneme relationships were used--a modification of Bennett's method of locating position of error in a word and Weber's Graphic Similarity Index.

At what position in a word do errors tend to occur? Judging by research cited above (Bennett, 1942; Marchbanks and Levin, 1965; Williams, Blumberg, and Williams, 1970; and Venezky, Chapman, and Calfee, 1971), one would expect more errors to occur in the medial and final positions than in the initial position. The studies do not agree as to whether middle or final position tends to attract more errors. In order to record error location in mistaken responses in the present study, a second measure of the grapheme-phoneme correspondence of response to stimulus was used--an adaptation of Bennett's method of

specifying the location of an error in a response. Every substitution error was rated according to the location of the error in the word in which it occured: initial phoneme, middle phonemes, or final phoneme. Note, for example, these positions in the word dog /dɔg/. Initial, medial, and final phonemes are /d/, /ɔ/, and /g/, respectively. In the word feather /fɛðɚ/ there are two phonemes in the medial position--/ɛ/ and /ð/. The initial phoneme is /f/, and the final phoneme, /ɚ/.

Weber's Graphic Similarity Index (1970b) is a recently devised means of determining degree of graphic similarity between printed stimulus and erroneous response. The more letters that match, the higher is the score. The same formula also takes into account similarity of word length. Furthermore, values are assigned to the various positions within the word where stimulus and response match; these number values are intended to parallel the importance of the different word positions as cues for recognition of words. More points are given, for example, for first letters matching in stimulus and response than for last letters matching. Identical shared pairs of letters also earn additional points. Weber determined the validity of her Index by having university students rank order pairs of randomly chosen stimulus words and erroneous responses in order of greatest approximation of stimulus and response. On the basis of Weber's findings from using the Index, it was expected that the better the reader, the higher would be the graphic similarity scores of his errors--when these readers of diverse abilities are compared at specific levels of proficiency. Furthermore,

the graphic similarity scores of all readers would presumably vary considerably according to the difficulty of the reading material. The writer hypothesized, partly on the basis of Biemiller's (1970) findings, that readers of all ability levels depend to a greater extent on contextual cues when the reading material is relatively easy. Dependence on graphic cues should increase as the vocabulary and syntactic difficulty of the reading material increases. Behavior in this regard is compared in the present study across proficiency groups.

Determining Use of Syntactic Context

There are three effective measures proposed for evaluating the use of syntactic context in this study. One of them, used by Biemiller (1970) and others, is simply an examination of the syntactic appropriateness of the error to the sentence through the point of, but not beyond, the error. This measure is based on the assumption that, contextually speaking, a child's response to a word generally depends on what preceeds it--the parts of the sentence already read. Furthermore, in his oral reading error study using children in grades two, four, and six, Allen (1969) found a high correlation between grammatical fit of error with preceeding context and with the sentence as a whole. Note the errors in the following sentences:

Stimulus: Here is the house.
Response 1: *Here is the he.
Response 2: Here is the horse.
Response 3: Here is the heavy....

The first response is clearly unacceptable; articles do not preceed pronouns in English. The second response is syntactically acceptable. While the third response would change the sentence structure, and most

likely the meaning as well, the response is acceptable since heavy can follow Here is the; one could complete the sentence beyond the point of error--Here is the heavy sledge-hammer, for example. It has been suggested that, at a given level of difficulty, children of all three proficiency groups should use context with almost equal proficiency; hence it is expected that there will be little divergence among the three groups in appropriateness of error to preceeding context--when comparisons are made at specific levels of error rate.

One means of determining the extent to which children use syntactic context is to analyze the patterns of errors within phrases. Recent psycholinguistic research has shown that decoding takes place in terms of phrases, the divisions of which are determined by immediate constituent analysis. Thus if phrases are the linguistic units for reading, one would expect an effect similar to that observed by Weber (1970b) in short sentences: more errors should occur at the beginning of phrases since the grammatical context should be least constraining at this point.

For many years linguists interested in syntax have used immediate constituent analysis to divide sentences into linguistic segments. For purposes of constituent analysis, linguists have identified two parts of the sentence--subject and predicate--as the basic syntactic units within the sentence. Thus in constituent analysis the first division is made between subject and predicate; these, then, are the first immediate constituents of a clause. In a complex or compound sentence, the first division is made between the clauses. Pike (1943) presented the first detailed procedure for constituent analysis; he set

up a six-layer scheme for breaking down sentences into immediate constituents, beginning with the clause and culminating in syllables, as he demonstrated in the diagram below. Division begins at Level I with the split between clauses in a sentence--if there is more than one clause in a particular sentence. The second-level division is made between noun phrase and verb phrase. At the third level a division is made between verb and direct object, and article and noun and modifiers. This dividing continues down to the divisions between syllables of a word.

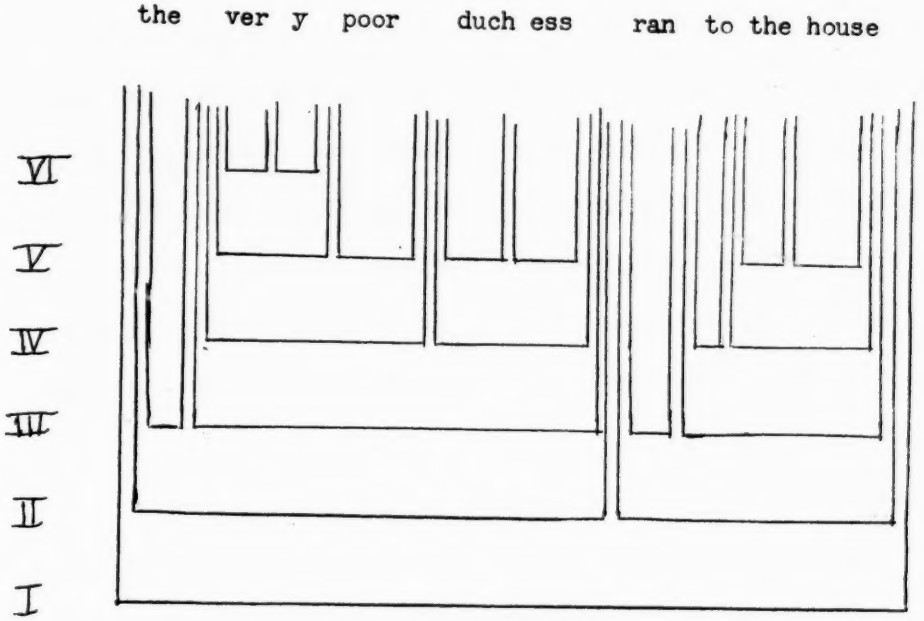

Figure 1. Pike's representation of constituent analysis.

Rather than a breakdown, one can consider the process, as Wells (1947) proposes, as a series of "expansions." He cites the sentence The king of England opened Parliament as an expansion of John worked. That is, John expands into the king of England, and worked

becomes opened Parliament. Both Pike and Wells concerned themselves with proving that the basic division comes between subject and predicate. No other division lends itself to this sort of analysis, as can be seen if one attempts to deal with clases not properly divided, such as The old farmer leaned / on the fence or The senator / from Maine won the nomination. In the former example sentence the sequence to the right of the division will be left not broken into the basic elements while the right-hand segment of the second sentence can be analyzed only as far as from Maine won since from Maine modifies senator.

Transformational grammarians--most notably, Chomsky (e.g., 1957, 1965)--have formalized the theory of constituent structure by studying the characteristics of the rules required to produce, or "generate," sentences with the appropriate constituent structure. Transformational grammarians use "tree" diagrams to depict the constituents of a sentence and their relationships to one another. One of the main components of a transformational grammar are the phrase structure rules, which are a means of accounting for the structure of sentences. The first phrase structure rule gives the components of the sentence (S)--noun phrase (NP) and verb phrase (VP):

$$S \longrightarrow NP \quad VP$$

The symbols on the right side of the "rewrite" arrow signify the constituents of whatever structure is indicated on the left. This diagram begins with the full sentence at the top and in descending order "divides" the sentence into its constituent parts, as is illustrated in the following tree diagram:

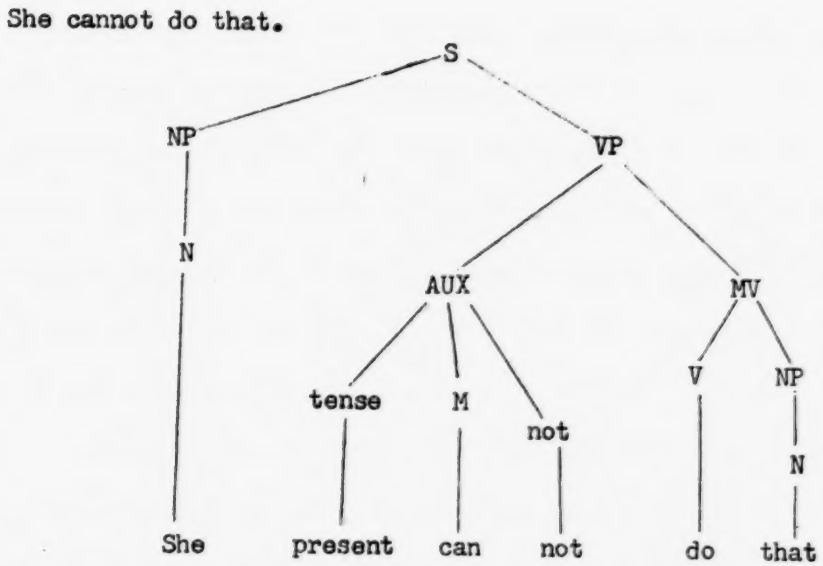

Figure 2. Tree diagram as used in transformational grammar. After the basic division of the sentence into two parts, each branch--noun phrase and verb phrase--is further divided into its respective constituents. Such a diagram makes the hierarchical ordering of the constituents more obvious.

Not only the reasoning of linguists but also experimenters in verbal learning have demonstrated quite convincingly that decoding takes place in terms of these units. Johnson (1965, 1968) found that subjects engaged in paired-associate sentence-learning tasks made more errors at the beginning of phrases than at other points in the phrase. The more major the phrase boundary, the greater was the tendency for errors to be made at that point. The boundary between noun phrase and verb phrase tended to attract more errors than any other point in the sentence, including other phrase boundaries. Studies of the eye-voice span (Levin and Turner, 1966; Schlesinger, 1968) have

also revealed that readers "chunk" the material by phrases. These two studies are similar in that the subjects' task was to read aloud, and at a predetermined point the experimenter would turn out the light and ask the subject to recite the words he had seen beyond what he had read orally. The last word in the "eye-voice span" in both studies tended to be the boundary of a major syntactic constituent. Other psycholinguists have done click placement studies on the assumption that in a perceptual task the boundaries of a perceptual unit--in this case a phrase--will attract interfering noises. That is, subjects will perceive a noise to have occurred in one of these natural breaks even when the noise may actually have occurred in the syllable immediately preceeding or following that break. In these studies the subject listens to spoken sentences through earphones. The experimenter introduces a click in one earphone, and the subject is asked where the click occurred in the sentence. Subjects in most of these studies--for example, Fodor and Bever (1965), Garrett, Bever, and Fodor (1966), Bever, Lackner, and Kirk (1969)--have tended to perceive interruptions at phrase boundaries, even when the click is actually located in the word or syllable immediately preceeding or following the boundary. This observation would seem to suggest that the subjects, and users of language generally, perceive phrase boundaries as divisions of the sentence and read in terms of these "chunks."

Thus this writer noted the number of erroneous responses that occurred at the beginning of major constituents--noun phrase (both subject and direct object), verb phrase, and prepositional phrase. The number of errors that occurred at the beginning of phrases and the

number of errors that occurred at other points in phrases were compared. If, in fact, readers do operate in terms of phrases, one would expect more errors at the beginning of phrases than elsewhere; syntactic and semantic context exerts the smallest effect at this point in the phrase.

Another measure of context usage is the position of the error in a sentence. Assuming a child uses grammatical context, Weber (1970b) reasons, there should be more errors committed at the point in the sentence where context has least effect--at the beginning. Weber did find that a high proportion of errors occurred at sentence boundaries, as she had predicted. Since sentences in more advanced reading materials tend to be longer and contain grammatical units that are not so closely related, it is likely that the reader reads by phrases rather than by sentences. However, an investigation of the tendency for errors to occur at the beginning of clauses might reveal the extent to which some readers read by units larger than the phrase--for example, the clause. The model of the reading process developed by Venezky and Calfee (1970) states that the reader proceeds in terms of "the largest manageable unit," a unit that will differ, presumably, according to the proficiency of the reader and the difficulty of the materials he is reading.

Furthermore, the characteristics of repeated sequences was noted; specifically, the extent to which repetitions of more than one word conform to phrases is examined. Since the purpose of a child's repeating more than one word apparently is an attempt to strengthen the context, syntactic and semantic, to deal with a troublesome word or sequence of words, it would seem that he should, when he repeats

sequences, regress to the beginning of a phrase--if, in fact, reading actually proceeds in terms of phrase structure.

In accordance with the suggestion in this chapter that readers of varying reading ability differ little in their ability to use context in passages where vocabulary does not present a problem, it was expected that there would be no differences between the three proficiency groups in this study in their tendency to make errors at the beginning of phrases and to repeat sequences that conform to phrase structure rules. Location of error in a phrase was examined only at the 0-4.9% level of error rate. Error location was not considered at the higher levels of error rate because as the reading material becomes more difficult, children make less use of syntactic context. Too, vocabulary difficulty would confound the results at higher error rates. The degree to which repeated sequences conformed to phrase structure, however, was observed at all levels of reading proficiency. Even when children are reading material that is very difficult for them, they should be able to perceive phrases--thanks to the numerous easily recognized function words in the language that denote these segments.

Finally, erroneous responses were classified according to whether they are function words--articles, prepositions, and auxiliaries, for example--or content words. K. S. Goodman (1966, 1967) suggests that the good reader selects the most significant cues and de-emphasizes those that are less significant. Because of the redundancy of the English language and his intuitive knowledge of syntax, the reader need not concentrate on every word. In reading, states Goodman, the proficient reader tends to make numerous errors on function words,

omitting, adding, or altering such forms as articles, prepostions, and auxiliaries--words which convey less meaning than nouns, verbs and adjectives. As the reading becomes more challenging, however, readers should become more accurate in calling these structures, probably because, with an increased number of strange words, they must attend more carefully to each structure in the sentence.

The reader should respond differently to function words and content words in context. Fries (1962) stressed that it is necessary for receivers of communication to respond to function words as grammatical "items" rather than, or perhaps in addition to, responding to them as lexically meaningful units. Unlike nouns and verbs, function words have no markers that signal their presence in a sentence. That is, the structure words the and a, for example, indicate to the reader that a noun is "coming"; in the same sense had, have, is, or was indicate the presence of a verb in the perfect tense--though these four markers may also serve as the main verb of the sentence. But no structure words precede and predict had, have, is, and was. It is possible to substitute non-words for content words, and find that most readers will yet perceive the "syntactic sense" of the sentence. But grammatical signals would be lost if function words were altered, and the reader's perception of the syntactic units of the sentence would be decreased. Gleason (1961) calls function words the "mortar" which joins the more meaningful words. Such words as a, and, and the are present in all English communication. Since function words pervade the English language, young readers have encountered them frequently,

first in speech and finally in reading. It would seem, then, that function words should cause little or no problem in decoding, even for poor readers. In fact, one would expect the most common of them to be among the first sight words acquired by the beginning reader.

One reason for the tendency of readers to make errors on function words is the fact that such words carry less lexical meaning than do content words--and in some cases practically no meaning at all. Gleason (1965: 186 f.) states that, while most function words carry some lexical meaning, their "semantic quality is overshadowed by the grammatical." Because of the very simplicty and lack of meaning of many function words, then, readers might be expected to "gloss over" them, responding to the grammatical signal and ignoring the graphic aspects of the word for purposes of meaning. This does not mean that the reader is unaware of function words--even of those that convey no lexical meaning. Rather, in reading, orally or silently, he is attracted to, and attends more carefully to, the more lexically meaningful words since presumably he is reading for meaning. Even young children probably realize that it is often possible to make lexical sense out of communications that contain no function words--for example, a telegram or a road sign.

A comparison was made in this study of the three proficiency groups in the proportion of short function words and short content words missed. On the basis of syntactic measures, it would be expected that there would be no differences in proportion of content word errors and function word errors. That is, in keeping with the argument advanced earlier that good, average, and poor readers do not differ in their proficiency

in using syntactic context, there should be no difference in the proportion of function words and content words missed at given levels of reading proficiency. However, from the hypothesis that better readers are superior to poorer readers in dealing with grapheme-phoneme relationships, it was expected that this prediction would not hold true. It seems likely that the poor reader must analyze every word carefully--even the relatively easy structure words--to perceive grammatical relationships as well as to break the code. Thus the poor reader seems likely to make proportionately fewer errors on function words than the good reader--not because he is unsure of the syntax but because he must decode even the easiest words in order to perceive grammatical relationships. In this regard, group differences in graphic similarity on function words was noted. The errors that the poor readers make on highly frequent words might bear a greater graphic and phonetic relationship to the printed word than the errors by good readers on the same types of words.

Summary of Hypotheses

In this study six hypotheses were tested. Three of them were directed to the readers' use of grapheme-phoneme correspondences, the other three to the readers' use of syntactic context.

The first hypothesis was addressed to the location of the error in mismatches:

> Hypothesis 1: The substitution errors of better readers show a greater tendency to match the printed word in both initial and final position than do the substitution errors of the poorer readers.

The better the reader, the more his errors reveal attention to the

graphic display. Concomitantly, the writer expected that the errors of the poor readers would tend to be more gross, with more responses that matched the stimulus in neither initial nor final position or in initial position only. It was recognized that the results of this hypothesis would likely be influenced somewhat by word length and word frequency, and attention was given these factors.

The second hypothesis was concerned with the reading process generally--the interaction of use of context and use of grapheme-phoneme relationships under different conditions of reading difficulty:

> Hypothesis 2: As reading material becomes more difficult, the substitution errors of all three proficiency groups show a greater graphic similarity to the printed word.

Graphic similarity was determined by using Weber's Graphic Similarity Index (1970b).

Hypothesis 3 made a prediction about the effectiveness of readers of the three proficiency groups in using grapheme-phoneme relationships at specified levels of error rate:

> Hypothesis 3: The better the reader, the higher is the graphic similarity of his errors to the printed stimuli at any given error rate.

Again, Weber's Index was used for determining graphic similarity. The graphic similarity scores of two subclasses of errors--responses that violate syntactic context and non-word responses--were examined separately. When an error is ungrammatical, the reader is apparently paying greater attention to the graphic display than to the context--as Weber (1970a) noted in her study of first-graders. Likewise, the writer thought that non-words might be found to be higher in graphic similarity to their printed stimuli. The circumstances in which children respond

with non-words were noted. It seems likely that the child who attends especially to the graphic display would prefer to create a graphically and phonetically "appropriate" structure rather than resort to uttering a legitimate word--meaningful or not. Perhaps it is a matter of context disintegrating to the point that the reader loses the meaning of the passage, cannot conceive of an appropriate response, and utters any combination of sounds at random.

In the three measures concerning use of syntactic context, no differences were predicted between the three proficiency groups--when the errors of all three groups were compared at the same levels of proficiency. Hypothesis 4 was addressed to grammatical appropriateness of mismatches:

> Hypothesis 4: Syntactic appropriateness of errors does not differentiate readers of different levels of reading proficiency.

The measure of syntactic appropriateness was whether or not the error fit the context from the beginning of the sentence through the point at which the mismatch occurred.

Hypothesis 5 predicted location of errors in phrases on the assumption that syntactic and semantic information is organized according to these units:

> Hypothesis 5: Tendency to read by phrases does not differentiate readers of different levels of reading proficiency.

The measure of reading by phrases was the location of substitution in phrases, an adaptation of a measure used by Johnson (e.g., 1965, 1968). This measure was used only in those passages in which a subject made errors on less than five per cent of the words read; in passages of

greater error rate this measure would have been confounded by performance on the word level. Errors were also examined with regard to position within clause to determine tendency to read by clause.

Another means of determining readers' awareness of phrase structure was the examination of repetitions of word sequences for conformity to phrase structure rules. Hypothesis 6 was based on the unproven assumption that repeated sequences do, in fact, conform to phrase structure rules:

> Hypothesis 6: Tendency to repeat word sequences by phrases does not differentiate readers of different levels of proficiency.

This hypothesis was tested by an examination of repeated sequences to determine to what degree they tended to be phrases. Intergroup differences in this regard would seem an indication of different degrees of awareness of or dependence upon syntactic cues. Repetitions were scrutinized at all levels of error rate; the writer believed that vocabulary difficulty affects a child's awareness of syntactic context only minimally.

In this chapter the problem chosen for the study was introduced, relevant literature cited, and hypotheses stated. The writer chose to examine the question of what differences in reading strategy, if any, exist between good, average, and poor readers. One possible way of examining reading strategy is through a linguistic analysis of errors made in oral reading. In this study the writer was particularly interested in differences in error patterns between good, average, and poor readers. The problem of whether oral reading errors differ according

to reading proficiency in quality as well as quantity was also of interest. It was believed that answers to these questions might reveal different reading strategies among the different types of readers. Six hypotheses pertinent to these questions were stated. In the following chapter the method of testing them is discussed in greater detail.

Chapter II

METHOD

The analysis of oral reading errors of good, average, and poor readers in the present study focused on an important question: to what extent do the erroneous responses of good, average, and poor readers approximate the stimulus words? The writer attempted to identify the differences between good, average, and poor readers at the third-grade level by examining the extent to which each type of reader uses contextual and graphic cues. The independent variable in the study was reading ability. The dependent variables were various measures of proficiency in using grapheme-phoneme correspondences and syntactic context.

The writer, then, wished to note the performance of readers varying in reading proficiency, keeping level of error rate constant. By doing so it was possible to observe the poor readers' use of syntactic context when word attack was not a problem. It was also possible to note the effect of difficulty of reading material on word attack behavior under different conditions of reading difficulty. Thus good, average, and poor readers were presented reading materials of various levels of reading difficulty to be read orally. Every subject, regard-

less of reading proficiency, began reading very easy materials--passages which even the poorest readers in the experiment were able to read proficiently. The subjects then proceeded through the readings that increased in difficulty to that point at which their error rate became excessive. The list of selections used is found in Appendix A.

Development of the Instrument

Since the experimenter wished to observe changes in reading strategy as the subjects advanced to more difficult reading materials, readings that were appropriately difficult at each grade level, one through nine, were chosen. All of the stories selected came from two sources of reading materials typically used in classrooms: the Science Research Associates (SRA) Reading Laboratory and the Madison (Wis.) Public Schools Informal Reading Inventory, a compilation of stories at various grade levels of difficulty. The selections and their sources are listed in Appendix A. Materials being used in the class instruction of the subjects were avoided.

There are basically three factors involved in controlling for the difficulty of reading material: (1) sentence structure, (2) vocabulary difficulty, and (3) what is sometimes called "concept load." Any one of these three factors can account for the difficulty of a passage. Only the first two factors were intended to be variables in the present study. Thus the experimenter, in choosing materials, tried to keep the effect of "concept load" constant by choosing materials in which the subject matter would be familiar to most children. Stories with simple plots were preferred. Materials at the upper grade levels were examined especially carefully to be sure that the selections chosen would

meet these criteria as well as possible at the third-grade level. Possible selections were evaluated for language difficulty according to three criteria: (1) publishers' ratings of the difficulty of their own materials, (2) readability formula scores, and (3) children's oral performance on the selections presented them in a pilot study. The readability formulas used were the Spache formula (1953) for grades one through three and the Dale-Chall formula (1948) for grades four and beyond. Both formulas take into account the average number of words per sentence as a measure of syntactic difficulty and the number of difficult words. The latter is determined in both formulas by counting the words in a passage that are not found in a specified word list. Since some good readers did not make a high error rate even on ninth-grade selections, passages rated grades ten, eleven, and twelve in difficulty were added. These materials for grades ten through twelve difficulty levels were evaluated for difficulty by publishers' ratings and readability formulas, but were not pilot tested.

The Pilot Study

The writer carried out a pilot study in December, 1970, in West School, Stoughton, Wisconsin. Eight third-graders participated in this phase of the study. They were grouped by reading proficiency according to teacher judgment: three good readers, two average, and three low. There were two reasons for doing the pilot study. First, the experimenter wished to make a final check on the levels of difficulty of the stimulus materials. Also, it seemed essential to do a "trial run" with children in order to make decisions on the details of the final procedure.

Each subject in the pilot study read orally the stories in the two alternate sets of stories, both of which had been evaluated for vocabulary and syntactic difficulty on the basis of publishers' ratings and readability formulas. Each of these two alternate sets of stories contained approximately three hundred words at each grade level, one through nine--a total of about 2700 words in each set of stories. Each subject read the passages to the examiner, beginning at the grade one level and proceeding to as high a level as he was willing until he made so many errors as to make the attempt seem pointless. These performances were tape-recorded.

The Final Instrument

From the two sets of stories used at each grade level, one was chosen for the final study. Errors tallied were substitutions, omissions, additions, and reversals, in conformity with standardized oral reading examinations and informal reading inventories. The number of errors on each story was totalled for each child. A mean error rate was derived by computing the average number of errors for children in each proficiency group. The passages for the final study were chosen so as to provide as fine and even a gradation in difficulty as possible from level to level. The mean error rate of each passage in both series for passages of grade one through grade five difficulty is charted in Table 1. The means of the passages chosen for the main study are indicated by an asterisk.

Table 1

Mean Error Rate in Two Sets of

Passages--All Subjects (N=8)

Difficulty of Material	Series 1	Series 2
Grade 1	24.4	13.8*
Grade 2	37.3	21.0*
Grade 3	26.6	34.3*
Grade 4	41.8*	35.7
Grade 5	48.5*	41.3

*chosen for the final study

Only the good readers proceeded beyond the fifth-grade level materials. Their mean error rates appear in Table 2:

Table 2

Mean Error Rate in Two Sets of

Passages--Good Readers Only (N=3)

Difficulty of Material	Series 1	Series 2
Grade 5	24.0*	21.6
Grade 6	23.8	30.3*
Grade 7	38.5*	32.0
Grade 8	40.7	42.4*
Grade 9	48.0	51.2*

*chosen for the final study

The readings at the first-grade level were very easy so that even the poorest readers had a chance to perform well on at least one set of materials. At the same time selections at the upper end of the difficulty range were made sufficiently challenging for even the best readers so that they achieved a high rate of error on at least some passages. Only then would it be possible to compare good, average, and poor readers at specific levels of proficiency. From such a comparison one might be able to determine the extent to which the poor reader's problem is one of use of syntactic context or one of inadequate mastery of grapheme-phoneme correspondences.

The materials chosen did meet the objective of providing some easy tasks for poor readers. However, as stated above, it was found that the ninth-grade level materials were not sufficiently difficult for a few of the good readers, so further passages were added to the collection, ranging as high in language difficulty as grade twelve. Again, the experimenter was careful to choose stories in which the subject matter itself should pose little or no problem for third-graders. Language difficulty was judged on the basis of publishers' ratings and readability formulas. As was noted above, these high-school level passages were not pilot-tested.

The selections collected for each grade level were then divided into an average of three passages, averaging one hundred words each for each grade level. This was done for two reasons: (1) to be able to present the selections to the subjects in three sessions rather than at one time and hence avoid tiring the subjects and (2) to retain the progression from easy material to difficult in each session so that no

subject would be jolted by resuming reading in materials beyond his depth. Thus most students covered one set of materials in about fifteen minutes. There were an average of one hundred words of readings at each grade level presented to the child in each of three sessions. At each session the child began reading material at the first grade level and proceeded to the point at which he made more than twenty errors on every one hundred words.

The Study

Subjects

Third-grade children were chosen as subjects for the study. The children were tested in February and March. The writer felt that children of this age might be appropriate subjects because by the third grade it is easier to delineate good, average, and poor readers than at earlier grade levels. By this time children, at least in the school system which cooperated in this study, have passed the beginning reading stage. One might also argue that by the third grade the poor readers are not as likely to have become thoroughly frustrated with learning to read--as might be the case with poor readers at higher grade levels.

Selected children in the Stoughton, Wisconsin, public schools participated as subjects in the experiment. Stoughton, a city of 5500 people, has a population blend of rural, small city, and suburban elements. It is located about fifteen miles east of Madison, the state capital and a city of some 170,000 residents. The reading method used in the school system has been a basal reader plan, supplemented by an extensive phonics program.

The thirty participants in the final study were randomly selected

from pools formed from all four third-grade classes at Yahara and West schools. All members of these classes were divided into three groups--good, average, and poor--according to reading proficiency. Reading proficiency for each child was determined by two measures: (1) performance on the comprehension and vocabulary subtests of the Gates-MacGinitie Reading Achievement Test (Level C, Form 2) and (2) teacher's estimate of each student's reading proficiency. Teachers made this rating by dividing the class into thirds according to reading proficiency. The degree of concord between test result and teacher judgment can be seen in Table 3:

Table 3

Agreement on Placement of Subjects

by Teacher and Test Results

Teacher	No. of Students in the Class	No. of Cases of Agreement	Per cent Agreement
01	25	17	68%
02	29	21	72%
03	31	19	61%
04	27	18	66%

For a child to be eligible for the random selection pool, it was necessary that both teacher judgment and test results indicate his placement in the same third of the class. Finally, in order to increase the homogeneity of the groups, the ten per cent of the eligible subjects who were nearest each of the two divisions between proficiency groups were dropped. It was believed that, if these reader proficiency groups had distinct characteristics in reading

behavior, they would more likely be revealed if the borderline subjects were dropped. None of the children who participated in the pilot study was eligible for the general study. After these initial procedures, the experimenter randomly selected subjects from the three groups, taking an equal number of girls and boys from each group. The composite score--vocabulary and comprehension--for each student is given in Table 4:

Table 4

Performance Scores of Subjects on the Gates-MacGinitie Reading Achievement Test (Level C, Form 2)

High Group		Middle Group		Low Group	
Subject	Raw Score	Subject	Raw Score	Subject	Raw Score
29	47.0	10	36.0	08	25.5
25	45.5	16	36.0	14	24.5
27	45.0	09	35.5	23	24.5
24	44.5	17	35.0	03	24.0
20	43.5	13	34.5	01	23.0
05	43.0	22	33.5	12	20.0
07	43.0	26	32.5	28	20.0
19	43.0	02	32.0	30	20.0
21	41.5	06	32.0	04	15.5
11	39.0	15	30.5	18	15.0
Mean score	43.5	Mean score	33.8	Mean score	21.2

Procedure

Each child read the graded passages aloud to the experimenter. The total reading task was divided into three sessions. To avoid the effect of fatigue on reading performance, each reading session was limited to a maximum of fifteen minutes. In each session the child began reading the first-grade selection and continued reading through the passage in

which his error rate exceeded twenty per cent. Error rate calculation conformed to the scoring procedures of oral reading tests generally. Substitutions, omissions, refusals or non-responses, additions, and reversals of words were counted as errors. Repeated errors were counted toward the error rate for a given passage only once.

From observations made in the pilot study, the writer decided that subjects should not be coached when they hesitate or refuse to utter a word. Pronouncing troublesome words for the child seemed an artificial means of bolstering the effect of context; it would likely affect one of the factors to be observed in this study--the effect of miscalls on the child's use of context.

Before each session the writer explained the nature of the task to the child, preparing him for the progression to difficult materials. The subject was told that he would not be coached. He was encouraged to guess at those words he did not recognize. He was assured that he could stop whenever he wished and that the session would last no longer than fifteen minutes, whatever his endurance. Every effort was made to keep the atmosphere relaxed.

During each subject's rendition, the writer sat at the side of, and slightly behind, the subject to keep from influencing the subject by any reactions. Some transcribing was done as the subject read. All performances were tape-recorded for later transcription of errors.

An examination of the children's performance in the pilot study showed that there was little reason for the subject's progressing beyond materials in which their error rate exceeded fifteen or twenty per cent. In each of the three sessions, then, the child read an average of one

hundred words at each grade level to the point at which his error rate disqualified him. When the subject reached the twenty per cent error level, he was asked to stop. It was intended that a subject complete one set of stories in each session, progressing from grade one materials through the passage in which the error rate exceeded twenty per cent. For those readers who were unable to complete a set of passages in the fifteen-minute time limit, the task was adapted somewhat at the following session. Rather than starting the child at the advanced point at which he had stopped at the end of the previous session, the experimenter provided a few minutes "warm-up" time; the child first read a few paragraphs from easier stories he had read earlier. Thus any apprehension that might result from resuming reading in advanced selections was avoided. These "warm-ups" were not transcribed, nor was any "warm-up" material ever reread by subjects at a later time for analysis.

Data Gathered

The graded passages were the basic unit of error analysis. There were three of these passages at each grade level, each of which averaged one hundred words. These units provided a means for comparing the performance of good, average, and poor readers. The rate of error was determined for each passage. Comparisons were made at the following levels of proficiency: (1) 0.0-4.9% error rate, (2) 5.0-9.9% error rate, (3) 10.0-14.9% error rate, (4) 15-19.9% error rate, and (5) 20% and above error rate.

Errors were transcribed phonetically from the tape recordings onto printed copies of each story. For purposes of this study, a substitu-

tion was defined as any error in calling a word, whether the alteration in the response changes pronunciation or meaning or both. A reversal was any transposition of two or more words. Additions were words added by the child that did not resemble the stimuli in graphic or phonetic form. Omissions were words in the text that were not rendered by the subject in any recognizable form. Refusals or non-responses are similar to, and sometimes difficult to distinguish from, omissions. In the present study a response was considered a refusal when, in the judgment of the experimenter, the child attempted to deal with the word--usually indicated by a lengthy pause--but made no audible response before proceeding to the next word. On the other hand, the apparent lack of a perceptible pause was generally taken as evidence of an omission. In the case of errors involving two or more wrong attempts, only the first attempts were analyzed in this study since it would seem that a child's first attempt is more indicative of his reading strategy.

Non-words presented a special problem, as children offered a substantial number of them when they read especially difficult material. It was decided to transcribe them phonetically. In comparing phonetic resemblance of stimuli and responses, then, the phonetic equivalents were used. In order to compare graphic similarity, however, as in Weber's Index, it was necessary to render the phonetic forms of these non-words in graphic form. Thus the experimenter assigned the graphic letters to non-words that are most commonly associated with the phonemes uttered. There were, of course, cases in which an utterance could be rendered in two different graphic representations. If, for

example, a child gave the response /fʌt/ for the word put, the word could be graphically transcribed futt or phutt. Where there were such alternate graphic equivalents for a given phoneme or combination of phonemes, the form that bore closest resemblance to the stimulus was chosen. The response /fʌt/, then would have been rendered graphically as phutt. The dictionary (Webster's New World Dictionary of the American Language--College Edition) was the determiner of whether or not a response was considered a legitimate English word. Erroneous responses to proper nouns were not considered non-words since most names are not meaningful words in the dictionary sense in the first place. Even a mispronunciation that closely resembled a legitimate English word--for example, ashamed pronounced as ashammed--is counted a non-word error; such a mispronunciation would seem to indicate that that particular stimulus word is not meaningful to the child. Mispronunciations arising from dialect or idiolect differences and common mispronunciations of words (for example, often with the t pronounced and the initial syllable of professional pronounced as per-) were counted as legitimate English words.

Analysis

In analyzing the data, there was first a comparison of proportions of error types among good, average, and poor readers at the same levels of proficiency. This consisted of a tabulation of errors. There were two measures of degree of similarity of stimulus and response on the word level--the Weber Graphic Similarity Index and measures of error location within a word. Use of syntactic context was measured in three

ways: syntactic appropriateness of the error to the sentence up to the point at which the error occurred, position of error in the clause, and conformity of repetitions to phrase structure rules.

Measures on the Word Level

Two means of specifying location of error in a word were used. In both cases the error was noted according to the position or positions in the word in which it occurred--initial, medial, or final. Of course, more than one of these positions could be faulty in any given error. For the first measure, phonological equivalence of the three positions of both stimulus and response were evaluated. Initial position was defined as the first phoneme of a word; final position was the final phoneme. All phonemes other than intial and final were considered medial position. Note the following word pairs as examples:

printed words	responses	position of error in words
light /laɪt/	fight /faɪt/	initial
feel /fil/	fail /fel/	medial
forge /fordz/	forget /fɚgɛt/	medial, final

The authority for all word pronunciations was A Pronouncing Dictionary of American English (1953) by Kenyon and Knott. Where regional alternatives appear in this dictionary, the Midwestern pronunciation was chosen as the standard for the printed word. Errors were designated in the following manner:

Type 1 Error--match in neither initial nor final position (dog-cat)
Type 2 Error--match in initial position only (dog-day)
Type 3 Error--match in final position only (dog-big)
Type 4 Error--match in both initial and final position (dog-dig)

Two-letter words, having only initial and final positions, were listed separately. Non-words were created phonetically by the experimenter

from the children's responses. It should be noted that in phonetic comparison, some errors involving the final -e are recorded as medial errors rather than as final errors, as one would expect on sight, for example: mat /mæt/, mate /met/. Also, the same graphemes may be rated phonetically different. For example, even /ivən/ and ever /ɛvər/ (a Type 1 Error) are graphically, but not phonetically, alike in initial position.

Substitution errors were divided into word groups by word length. All errors on words that contained less than three phonemes--that is, words without a medial position, phonetically speaking--were omitted from this part of the analysis. The remaining stimulus words were distributed into two categories--short words (three or four phonemes in length) and long words (five or more phonemes). The six most frequently occurring words in the language--a, the, and, to, of, and in--were omitted from this analysis. In addition, long words were evaluated for similarity in the second letter and in the second letter from the end of stimulus and response.

In addition, short words were classified as function words and content words. There are two listings of function words in English known to this writer--those by Fries (1952) and Francis (1958). The two lists are in close agreement, though the methods of presentation differ. Fries provides frames; words which fit into the slots in these frames are considered function words. Francis uses labels (prepositions, determiners, and the like), with lists of words for each category. The experimenter has chosen Fries's listing as the guide in this study,

largely because there are words that can be function words and content words under different conditions, making the word list approach seem inappropriate. In the sentences below, for example, have and there are function words in the first two sentences, but content words in sentences 3 and 4:

1. There is more than one way to skin a cat.
2. The police have arrested Joe's grandmother.
3. There is my car.
4. Ev and El have a canoe.

Thus the word list approach seems less reliable. The frame format allows for the fact that the same word can be classified as a function word and as a content word, depending upon its context.

The experimenter also did further phonological analysis on mismatches in a special group of long words. In those long words in which stimulus-response matched in initial and final position, notation was made as to whether or not the second phoneme from the beginning and the second phoneme from the end matched. In the pair flight /flaɪt/ and flint /flɪnt/, for example, the writer would note that the second phoneme--/ l /--is identical in the printed word and in the child's response, but that the second from the last phonemes differ--the diphthong /aɪ/ in the printed stimulus and the /n / in the oral response.

The Weber Index was applied to all substitutions in which there were graphic differences between the printed word and the erroneous response in order to determine degree of graphic similarity. Mean graphic similarity scores were computed for each child and for each proficiency group at all levels of error rate.

The Graphic Similarity Index, described in Chapter I, is a recently

devised means of determining degree of graphic similarity between stimulus and response. Weber (1970b) gives it in the following formula:

Table 5

Graphic Similarity Index Formula

$$GS = 10\left[\frac{(50F + 30V + 10C)}{A} + 5T + 27B + 18E\right]$$

F = the number of adjacent letters in the same order shared by P (printed word) and R (response)
V = the number of pairs of adjacent letters in reverse order shared by P and R
C = the number of single letters shared by P and R
A = average number of letters in P and R (total number of letters in P and R divided by 2)
T = ratio of number of letters in the shorter word to the number of letters in the longer word
B = 1 if the first letter in the response is the same as the first letter in the printed word; otherwise B = 0
E = 1 if the last letter in the response is the same as the last letter in the printed word; otherwise E = 0

A pair of words—for example, show-snow—would be evaluated in the following way:

$$GS = 10\ \frac{50(1) + 10(3)}{4} + 5\left(\frac{1}{1}\right) + 27(1) + 18(1)$$

The GS for show-snow would be 700. The following are a few more examples of possible errors and the GS scores derived:

shake/shape	780
and/also	347
bring/birth	440
rain/rained	683

Numbers were assigned to each possible position within the word according to how powerful a cue each position is in word recognition.

Separate scores were tallied for some of the special types of substitutions—non-words and responses that violated syntactic context.

The first two special classes were analyzed separately in order to note whether children, when they gave such responses, were depending to a greater extent on the grapheme-phoneme relationship than usual. If this was the case, one would expect a higher graphic similarity score on these mismatches. Suffix errors were separated for scoring to determine their effect on mean graphic similarity scores. It was believed that, since suffix errors generally result in little alteration of the printed word, a preponderance of errors involving only suffixes might inflate the mean graphic similarity score of any child or group.

Substitutions involving certain very common function words--the, of, and, to, a, and in--were omitted from graphic similarity analysis. According to the Kučera and Francis (1967) frequency count[4] of the most common words in the English langauge, these are ranked, in the order given, as the six most often used words. The dividing line for purposes of this study was made between in and that, the latter being the seventh word in the list, because of the considerable gap in the frequencies of the two words as recorded by Kučera and Francis; in occurred 21,341 times in the corpus while that occurred 10,595 times. Inclusion of errors on these words might have unduly deflated the scores of those students who habitually substitute a for the, for example. Since these words are so common, it is assumed that even the poorest readers in this study would recognize them in print and that erroneous responses could

[4] The Kučera-Francis list was chosen over the more commonly used Lorge-Thorndike Teacher's Book of 30,000 Words (1944) largely because of its recency. Other advantages of the Kučera-Francis compilation, as well as weaknesses of the Lorge-Thorndike list, are discussed in detail by Shapiro and Gordon (1971).

be attributed to the influence of context. Thus this sort of error was more appropriately treated in other sections of this study—in the tabulation of types of errors and in the comparison of rate of error on function words as compared with errors on content words.

Measures on the Syntactic Level

Appropriateness of the error to the syntactic context in which it is located was noted. As was stated in Chapter I, in this study an error was considered grammatically appropriate if it fit syntactically into the sentence up to the point of error, regardless of whether or not stimulus and response were of the same form class. Percentages of errors that violated syntax were tabulated for each child and then for each proficiency group on each passage and at the various levels of error rate.

A differentiation was made between errors according to their locations in phrases. Sentences in the reading materials were divided into phrases in accordance with constituent analysis, as described in detail in Chapter I. Some sentences that contained discontinuous constituents were omitted from this analysis. Only the first word of a constituent was considered in this analysis. Each unit was not, of course, of equal length in number of words, as has been true in verbal learning studies generally. But the fact that the phrases in the stimulus materials vary in length and overlap should not nullify the tendency for readers to operate in terms of phrase structure rules. Schlesinger (1969), in studying the reader's use of constituent structure in decoding, found that readers "chunked" material into units of varying sizes—but that these units, regardless of size, tended

to correspond to phrase structure. In the present study, then, the fact that constituents are of varying length should not have changed the fact that readers decode by phrases. It was expected that, in the oral reading situation, more errors would occur at the beginning of phrases than elsewhere in the phrase--provided the reader was working with easy material. This measure was limited to those passages in which the subject made no more than five errors per hundred words. In most sentences of the stimulus materials, then, the initial word of each phrase--noun phrase, verb phrase, and prepositional phrase--was identified. The number of errors that occur at the beginning of these phrases and the number of errors that occur elsewhere in the phrases were tallied. The results were compared across the three groups of subjects.

Furthermore, the position of all errors in the cluases in which they occurred was examined in all passages in which subjects make no more than five errors per hundred words. Clauses are, of course, composed of phrases. In their model of the reading process, Venezky and Calfee suggest that the reader operates by the "largest manageable unit" that he can "chunk" efficiently, be it a clause, phrase, or only a single word. Thus it seems plausible that some subjects, presumably the more proficient readers, would read by sentences. The work of Schlesinger (1968) and Levin and Turner (1966) indicates that the better the reader, the longer is his eye-voice span. An examination of the ratio of number of errors at the beginning of clauses and the number of errors elsewhere in the clauses should indicate the extent to which good, average, and poor readers operate in terms of clauses. Since the "largest manageable unit" should be longer for good readers,

good readers should make the most errors at the beginnings of clauses and the fewest errors at the end of clauses, for the beginning of a clause is the point at which contextual constraint is minimal. In this study, positions in clauses were determined by dividing the sentences into three sections according to the total number of words in the in the sentence. Every effort was made to make the sections as equal in number of words as possible so that in an entire story there was nearly always an equal number of words in each of the three positions within the sentence. Such linguistic units as phrases were not considered in dividing these sentences, nor was there any concern for surface structure-deep structure differences. Clauses of less than three words were excluded from this analysis unless they were preceeded by, or followed by, another clause in the same sentence. In that case, these two-word clauses were counted as part of the preceeding or following clause in the same sentence. Sentences containing "discontinuous constituents"--relative clauses or other interrupting elements--were simply excluded from this analysis. Clauses having two verb phrases, both three words or more in length, were further separated for analysis between the two verb phrases. Note the following sentence, for example:

> The jumping fireworks salesman cursed the careless lamplighter and called for the fire brigade.

This sentence consists of one clause with two verb phrases, cursed... lamplighter and called...brigade. The second verb phrase describes an action separate from that of the first verb phrase and thus brings a new, though related, set of meanings into play. That is, while the context keeps "building up" from the beginning of the clause through the word lamplighter, that line of meaning stops when the reader comes to the

second verb, called. Beginning with the verb called, a new line of contextual meaning comes into play. In terms of the present study, it was assumed that the beginning of the verb phrase called...brigade would attract more errors than would the end of that verb phrase—in the same sense that errors are more likely to occur at the beginning of a clause than at the end of it. Lengthy participial phrases were separated from the rest of the clause for the same reason, which might be made clear in the following example:

> Noting the footprints on the ceiling, Cedric concluded that his weird relatives had called.

Certainly the participial phrase Noting...ceiling is related to the rest of the sentence, but, again, two different actions are portrayed. It is suggested that context builds from noting through ceiling; with Cedric begins a new line of meaning, which builds through the rest of the sentence. Thus the writer deemed it necessary to further divide clauses containing two verb phrases and to treat participial phrases separately.

The mean scores for location of error in the clause were compared at the 0.4.9% error rate to determine the extent to which readers of different levels of proficiency draw upon context under different difficulty conditions. This measure is similar to one of Weber's (1970b) measures of first-graders' use of context. The tendency for errors to occur at the beginning of clauses can be observed only when the reader is working with relatively easy selections. At higher levels of reading difficulty the measure is confounded by vocabulary difficulty. Thus this measure was applied only to those passages in which the error rate was less than five per cent.

In Chapter I it was suggested that repeated sequences of words might also demonstrate chunking by phrases. That is, if readers actually do tend to operate in terms of phrases, they could be expected to go back to the beginning of a phrase when they repeat two or more words. This measure was applied at all levels of proficiency. The number of repetitions that constitute regression to the beginning of a phrase was compared with the number of repetitions that do not.

Repetitions involving discontinuous constituents were omitted from this analysis. Only those repetitions in which the reader could have regressed to words that were not at the beginning of phrases were considered. In the following repetitions, for example, the reader, if he would have regressed, would have had no choice but to begin the repeated sequence with a phrase:

Joe went to the circus.

Joe is the beginning of the noun phrase; in fact, it is the noun phrase. In the same way, went is the verb phrase, and to begins the prepositional phrase. To avoid biasing the results in the hypothesized direction, repetitions such as the one above were not tallied. Compare the above example with the repeated sequence in the following sentence:

That girl in the swim suit left early.

The regression from suit to the would violate constituent structure in that the reader does not go back to the beginning of a phrase--the word in. Similarly, backtracking from suit to girl would violate constituent structure in that the reader would cross a phrase boundary, regressing to a word that was not at the beginning of a phrase. On the other

hand, a return from suit to in would be a repetition compatible with constituent structure since in begins a prepositional phrase. Repetitions which cross phrases can be in conformity with constituent structure--provided the reader reverts to the beginning of a previous phrase that is "higher up the tree" (See diagram on Page 30).

SUMMARY

In this chapter the methods of testing the hypotheses have been described. A linguistic analysis of oral reading errors was the means used to test these hypotheses. Thirty third-grade children, chosen and grouped on the basis of reading proficiency, served as subjects in the study. Each of the three proficiency groups was composed of ten children, equally divided by sex. Analysis of the data was done by various measures of use of grapheme-phoneme correspondences and of syntactic appropriateness of errors.

Chapter III

RESULTS AND DISCUSSION

The six hypotheses stated in Chapter I follow:

> Hypothesis 1: The substitution errors of better readers show a greater tendency to match the printed word in both initial and final position (Type 4 Error) than do the substitution errors of the poorer readers.
>
> Hypothesis 2: As reading material becomes more difficult, the substitution errors of all three proficiency groups show a greater graphic similarity to the printed word.
>
> Hypothesis 3: The better the reader, the higher is the graphic similarity of his errors to the printed stimuli at any given level of error rate.
>
> Hypothesis 4: Syntactic appropriateness of errors does not differentiate readers of different levels of reading proficiency.
>
> Hypothesis 5: Tendency to read by phrases does not differentiate readers of different levels of reading proficiency.
>
> Hypothesis 6: Tendency to repeat word sequences by phrases does not differentiate readers of different levels of reading proficiency.

Results

Hypothesis 1

Long word errors were in accord with the hypothesis, but short word errors were not. The patterns observed in short word errors were

unexpected. In order to more effectively present and explain these patterns, long word errors and short word errors are considered separately throughout this discussion of Hypothesis 1.

Within both long word and short word categories the number of errors of each error type were tallied for the subjects in each group. Errors were classified according to the following categories:

> Type 1 Error--match in neither initial nor final position (dog-cat)
> Type 2 Error--match in initial position only (dog-day)
> Type 3-Error--match in final position only (dog-big)
> Type 4 Error--match in both initial and final positions (dog-dig)

In addition, the examination of matching in the second phoneme from the beginning of a long word and the second phoneme from the end of a long word necessitated the addition of subgroups under Error Type 2 and Error Type 4:

> Type 2' Error--match in at least first two phonemes and in no more than one phoneme at the end of the word (break-brought, pride-proud).
>
> Type 4' Error--match in at least the first two phonemes and in at least the last two phonemes (stamp-stump).

The raw data for individual subjects and for the three proficiency groups are presented in Appendix B. Errors were examined at five levels of error rate: (1) 0-4.9%, (2) 5.0-9.9%, (3) 10-14.9%, (4) 15-19.9%, and (5) 20% and above.

The heterogeneity G-test, or log likelihood ratio test (Sokal and Rohlf, 1969), was used to determine the extent of intergroup differences in proportions among the four error type patterns for short-word errors and long-word errors. The G-test is used with frequency data to deter-

mine if the differences in proportions within categories are significant. The formula for the G-test is as follows:

$$G = 2\sum^{a} f_i \ln\left(\frac{f_i}{\hat{f}_i}\right)$$

a is the number of categories involved in the comparison. The G-test is numerically similar to the chi-square, with both tests having similar distributions. When subgroups are involved, the G-test is computationally more feasible than chi-square.

No significant differences in percentages among the error types were found between the five rates in either short words or long words. So performances under all five error rates were combined within the long word and short word groups. The percentage of long word errors and short word errors are given in Table 6 while Tables 7, 8 and 9 show the intergroup differences in error type. A test of analysis of

TABLE 6

Percentages of Long- and Short- Word Errors

Proficiency Groups	Short-word Errors	Long-word Errors	Total
Good Readers	19.3%	80.7%	100.0%
Average Readers	24.3%	76.7%	100.0%
Poor Readers	26.0%	74.0%	100.0%

variance revealed no significant differences between proficiency groups.

Data on long word errors support the hypothesis. A percentage was computed of each type of error to the total number of errors in each proficiency group. Table 7 shows the percentage of each error type of long word errors that were made in each proficiency group.

TABLE 7

Percentages of Error Types--Long Words

Proficiency Groups	Error Types			
	Type 1	Type 2	Type 3	Type 4
Good Readers	4.5%	32.5%	9.5%	53.4%
Average Readers	7.7%	37.2%	10.9%	44.2%
Poor Readers	8.5%	39.5%	9.7%	42.3%
Significances				
Good-Average	G=11.2, $p<.001$	NS	NS	G= 8.7, $p<.01$
Good-Poor	G=13.8, $p<.001$	NS	NS	G=10.8, $p<.01$
Average-Poor	NS	NS	NS	NS

The raw data is contained in Table 21 in Appendix B. The percentage of error types in each group was compared by the G-test, as described above. Statistical differences between the combinations of the groups are found at the bottom of the table. Observations of long word errors support the hypothesis in error types 1 and 4. For Type 1 Error the difference is significant between good and average readers (G=11.2, $p<.001$) and between good and poor readers (G=13.8, $p<.001$). The difference between average and poor readers was not significant. For Type 4 Error the difference is significant between good and average readers (G=8.7, $p<.01$) and between good and poor readers (G=10.8, $p<.01$). The difference between average readers and poor readers did not reach signifi-

cance in Type 4 errors. There were no significant differences between the proficiency groups for Type 2 errors and Type 3 errors. It would indeed seem that the better the reader, the more attention he is likely to give to the final portion of long words. But the fact that the differences between average and poor readers were not significant would seem to indicate that the data demonstrates not so much a weakness of poor readers as a "strength" that distinguishes the good readers.

A rank ordering of error types by each subject within each group was done as another means of considering the relative strength of cue positions within words by groups. As can be seen by examining the raw data by subject (Appendix B, Table 21), the most frequently made error type for all of the good readers was the Type 4 Error. The Type 4 Error was the most common for six of the average readers; one average reader made an equal number of Type 2 and Type 4 errors. Five of the poor readers made more Type 4 errors than any other type. This, too, would seem to indicate a greater tendency among the good readers to attend more carefully to both beginning and end of long words.

Stimulus and response pairs of long words were also scrutinized for matching of initial two phonemes and final two phonemes. This was done in order to determine the extent to which readers of the three groups attend to the interiors of words. The percentages of each type of error--Type 2' and Type 4'--of the total number of long word errors is summarized for each proficiency group in Table 8, with significant differences indicated. The raw data is presented in Appendix B, Table 22.

TABLE 8

Percentages of Error Types 2' and 4'

to Total Long Word Errors

Proficiency Groups	Type 2'	Type 4'
Good Readers	38.6%	16.3%
Average Readers	36.4%	10.4%
Poor Readers	34.4%	7.6%
Significances		
Good-Average	NS	G=16.2, $p<.001$
Good-Poor	NS	G=34.2, $p<.0001$
Average-Poor	NS	G= 5.5, $p<.05$

The percentage of Type 2' errors to the total number of long word errors does not differ significantly between the three groups. The percentage of Type 4' errors to the total number of long word errors is significant between good and average readers (G=16.2, $p<.001$), between good and poor readers (G=34.2, $p<.0001$), and between average and poor readers (G=5.5, $p<.05$). These results seem to provide further support for the hypothesis that good readers attend more carefully to the endings of words than average and poor readers. The results also seem to indicate that the good readers are more accurate in their processing of the internal letters located near the ends of words.

Differences between average and poor readers on long words were not as great as one might expect, given Bennett's (1942) observations.

Several factors might account for this. First, Bennett's subjects, whom she described as "non-readers", may have been even poorer readers than those who participated in the present study. Secondly, it is implied that her subjects were taught reading in a program that emphasized teaching words as wholes. In a program in which learning whole words was emphasized rather than acquiring phonics skills, one might expect less transfer of reading skills to strange words--and hence a larger number of gross errors. Perhaps a strong phonics program is especially helpful for children with reading difficulties--at least with the decoding aspect of the reading process. Certainly a study similar in design to this one but with method of teaching reading as a variable might yield interesting information on this question.

Observations of errors on short words did not follow the overall tendency. In fact, the tendency for short word errors was the opposite of that of long word errors. The percentages of error types on short words between the proficiency groups is summarized in Table 9; the raw data is contained in Table 23 of Appendix B. Average and poor readers were more inclined to make errors indicating close attention to the graphic display while the good readers made a smaller percentage of graphically appropriate errors. The G-test indicated some significant differences between groups, but not in accord with the hypothesis. The good readers made significantly more Type 1 errors than did average readers ($G=17.5$, $p<.001$) and poor readers ($G=25.3$, $p<.001$). Under Type 4 errors, the good readers made significantly fewer errors than average readers ($G=13.5$, $p<.001$) and poor readers ($G=28.5$, $p<.001$).

The difference between average and poor readers was also significant ($G=4.7$, $p<.05$).

TABLE 9

Percentages of Error Types--Short Words

Proficiency Groups	Error Types			
	Type 1	Type 2	Type 3	Type 4
Good Readers	38.5%	29.3%	20.2%	12.0%
Average Readers	20.6%	38.9%	16.2%	24.3%
Poor Readers	16.7%	34.2%	16.7%	32.8%
Significances				
Good-Average	$G=17.5$, $p<.001$	NS	NS	$G=13.5$, $p<.01$
Good-Poor	$G=25.3$, $p<.001$	NS	NS	$G=28.5$, $p<.001$
Average-Poor	NS	NS	NS	$G=4.7$, $p<.05$

Two possible explanations for the difference in short word errors and long word errors came to mind. First, it seemed that a difference in the percentage of function words and content words might account for this observation. Function words--particularly short function words--are generally very high frequency words (Kučera and Francis, 1967), and, because the reader encounters these words so frequently, he is more likely to learn them as wholes and treat them as sight words. The nature of errors on short words might, then, be attributed to a difference in size of sight vocabulary. That is, good readers would seem more likely to deal with function words without analyzing them while poor readers might have to analyze even these very common structures. Secondly, it was thought that readers might attack short function words differently from the way in which they attack short

content words. A difference in the ratios of short function words and short content words might account for the discrepancy in proportions of error types between short words and long words.

Short word errors were broken down into two groups--short function word errors and short content word errors. As Table 10 indicates, there were differences in the proportions between the proficiency groups; the better the reader, the more likely he was to err on function words. While the good readers made a little over one-half of their errors on

TABLE 10

Percentage of Function Word Errors and

Content Word Errors--Short Words

Proficiency Groups	% Function Word Errors	% Content Word Errors	Total
Good Readers	55.4%	44.5%	100.0%
Average Readers	42.0%	58.1%	100.0%
Poor Readers	31.6%	68.4%	100.0%

function words, the poor readers made less than one-third of their errors on function words. So it seemed that the different proportions of function word and content word errors might account for differences in error type--if, in fact, word frequency and function is a factor in word attack.

Scrutinizing short word errors by error type revealed that subjects apparently do approach content words differently from function words, as the data in Table 11 seem to indicate. While over one-third of the errors on short content words are Type 4 errors, an insignificant

number of Type 4 errors occurred on function words. Conversely, readers were prone to make more Type 1 errors on function words than

TABLE 11

Percentage of Error Types on Short Word Errors:

Function Words--Content Words

	Type 1	Type 2	Type 3	Type 4
Function Words	37.9%	34.4%	26.1%	2.6%
Content Words	11.4%	36.4%	13.7%	37.8%

on content words. The ways in which subjects processed function words and content words is markedly different, indicating the importance of word frequency. It would seem that on function words readers give little attention to the graphic display, attending more to meaning and syntax than to the features of the word. Less frequent content words, however, despite their short length, need to be processed with somewhat greater care--perhaps because they are less likely to have been committed to memory.

A breakdown of short function word error types by reading proficiency group showed few differences between the groups in this regard. The good readers tended to make more errors in which there was no graphic match in initial or final position (Type 1 Error). None of the three proficiency groups made a sizable number of Type 4 errors--an apparent indication that subjects in none of the three proficiency groups were inclined to attend carefully to the graphic display when dealing with function words. The raw data for error

TABLE 12

Percentage of Error Types on

Short Function Words

Proficiency Groups	Type 1	Type 2	Type 3	Type 4
Good Readers	44.9%	31.1%	22.4%	1.6%
Average Readers	35.1%	34.9%	26.9%	3.2%
Poor Readers	33.7%	37.3%	30.0%	2.9%

types on function words is given in Appendix B, Table 24.

There is a marked difference in the percentage of error types on short content words and on short function words, as can be noted in Table 13. The better readers registered more Type 1 errors and fewer

TABLE 13

Percentage of Error Types on

Short Content Words

Proficiency Groups	Type 1	Type 2	Type 3	Type 4
Good Readers	15.0%	36.7%	17.0%	31.3%
Average Readers	11.4%	39.3%	11.1%	36.2%
Poor Readers	7.8%	33.1%	13.1%	45.8%

Type 4 errors than did the poorer readers. This trend between the proficiency groups is obviously the opposite of the trend observed with long words.

This reversal can probably be accounted for by the explanation suggested with short function word errors; the good reader was inclined to substitute short content words on the basis of meaning or syntactic appropriateness while the poor reader might have analyzed even short words

with considerable care. In fact, a comparison of short content-word errors and long word errors revealed that length of word had relatively little effect on the reading strategy of the poor reader. For example, the Type 4 Error--an indication of careful attention to the graphic features of the word--occurs in 45.8% of the short content word errors and in 42.6% of the long word errors made by the poor readers. While one might be tempted to conclude from this finding that poor readers are able to attack short words with relative ease and accuracy, two points should be kept in mind. First, as was noted in Table 10, poor readers made proportionately more function word errors than content word errors. Furthermore, Type 4 short content word errors comprise only a small proportion--12.8%--of the short word errors made by good readers while this type of error was made in 32.7% of the short word errors made by the poor readers. This could be indicative of a deficiency pattern in the poor readers--a weakness in processing the "interiors" of words. Since most of the short content words that drew errors in this study were of the consonant-vowel-consonant (c-v-c) structure, this preponderance of Type 4 errors might be an indication that poor readers experience greater difficulty in processing vowels than do good readers. In fact, more errors involving vowels were made on short words by poor readers (74.3%) than by good readers (56.7%). Thus the position of the mismatch in a word is likely influenced by factors other than word position--the type of sounds involved, for example.

But the problem of dealing with the middles of words should not be considered the result only of poor decoding skills. Rather, it

may be the result of the poor reader's failure to respond adequately to all of the reading cues. Dr. Robin Chapman suggested that the good reader in this study, responding more fully to syntactic and semantic cues, might more easily narrow the choice of response to the right word on the basis of only the first and last letters of a word. For example, when a child looks at a trigram that begins with c and ends with t, sentence structure, meaning, and graphic cues--in this case, the letters c and t--delimit the number of possible choices (cat, cot, cut) to the appropriate one. Thus a reader who responds effectively to the different cues would be less likely to make a Type 4 error on short words.[5] Perhaps the poor reader's apparent preoccupation with the graphemes limits his ability to respond to the other cues.

The observations summarized in this chapter might seem to conflict somewhat with the findings of Venezky, Chapman, and Calfee (1970), summarized in Chapter I. In the present study the tendency for readers to make more graphically appropriate responses--errors in which both initial and final positions match (Type 4 Error)--differed by proficiency group according to word length. On longer words, the better the reader, the more likely he was to make Type 4 errors. But errors made on short content words showed the opposite tendency across proficiency groups; the responses of the poor readers were more likely to match the stimulus word in the final position than were the errors of

[5] Personal communication, Dr. Robin Chapman, Assistant Professor of Communicative Disorders, University of Wisconsin.

good readers. Venezky, Chapman, and Calfee found at the lower grade levels that subjects in the upper quartile in reading proficiency were less likely to make errors in the final position of trigrams than were poor readers. This apparent conflict can likely be attributed to the fact that synthetic words in isolation were used as stimuli in the earlier study while real English words in context constituted the reading materials in this study. The word frequency factor--found to be significant in word attack strategy--can hardly be dealt with in a study in which synthetic words are used. It is probable that the subjects were familiar with most of the short content words in this study. These readers would not be familiar with the short synthetic words; hence they would likely approach these words in much the same way as they would approach strange English words, tending to scrutinize the graphic display. Also, as discussed above, the character of the sound involved, as well as the position of that sound, is a factor in word attack in this study.

Thus Hypothesis 1 is accepted, but with some qualification. While the hypothesis is supported in the analysis of errors on long words--words of five or more letters--the observations on short word errors seem to contradict the hypothesis. These differences have raised other points about the reading process and about the deficiencies of poor readers. Readers of all levels of ability displayed curiously different behaviors in responding to function words and content words. Of especial interest is the inclination of the poor readers to make more graphically accurate mismatches on short content words; this should probably be taken as an indication of deficits in word analysis,

sight vocabulary, and response to all of the cues involved in effective reading. The division of words into function words and content words might have been too gross a distinction for the types of analyses that were made. A hierarchy of divisions based on word frequency might be used more profitably in future research.

Hypotheses 2 and 3

> Hypothesis 2: As reading material becomes more difficult, the substitution errors of all three proficiency groups show a greater graphic similarity to the printed word.
>
> Hypothesis 3: The better the reader, the higher is the graphic similarity of his errors to the printed stimuli at any given level of error rate.

Hypotheses 2 and 3 are treated together because the same statistical procedure was used in treating both.

Mean graphic similarity scores were computed for each subject. The data at the fifth rate was omitted from analysis because of the large number of approximations that would have been needed; many of the good readers did not place at this rate. For each of the other four error rates a mean GS score was computed for each proficiency group. The mean scores for each group are found in Table 14. A two-way test of analysis of variance (Winer, 1962) was performed with repeated measures; this is summarized in Table 15.

TABLE 14

Mean Graphic Similarity Scores

Proficiency Groups	Error Rates			
	Rate 1	Rate 2	Rate 3	Rate 4
Good Readers	539.3	620.7	637.5	629.0
Average Readers	536.1	557.1	600.7	600.2
Poor Readers	428.5	568.2	580.8	589.8

As is indicated in Table 15, the effect of error rate, pertinent to Hypothesis 2, is significant ($p<.05$). As reading material becomes

TABLE 15

Overall Analysis of Variance Based on Mean Graphic Similarity Scores

Source	df	MS	F	P
Between Subjects:				
Proficiency Group	2	26,942.861	.747	NS
Subjects within Groups	27	36,054.915		
Within Subjects:				
Error Rates	3	69,186.732	3.316	.05
Error Rate by Proficiency Group	6	53,484.321	2.564	.05
Error Rate by Subjects within Groups	81	20,863.712		

more difficult, mean graphic similarity scores increase--a sign that readers of all proficiency groups attended more carefully to the graphic display as they encountered more difficult words. Thus the data support Hypothesis 2.

The reader proficiency variable (Hypothesis 3) failed to register statistical significance, although the results run in the hypothesized direction. Differences by reader proficiency were not significant. It can also be noted in Table 15 that the interaction between error rate and reader proficiency was significant ($p<.05$). Individual tests made at each error rate revealed significance at none of the error rates between the proficiency groups.

While the results for Hypothesis 3 were not expected, these findings are not surprising in view of the observations of the data for Hypothesis 1. Under Hypothesis 1 it was found that word length and word function have a considerable effect upon decoding behavior. Weber (1970b)

did not provide for these factors in her Index. Any meaningful use of the Index at the third grade level would seem to demand a revision of the Index to account for these factors. Too, if it is true that as children mature they tend toward responses more consonant with the syntax and the meaning of the passage than with the graphic display, this would presumably be reflected in the GS score.

Hypothesis 4

> Hypothesis 4: Syntactic appropriateness of errors does not differentiate readers of different levels of reading proficiency.

The proportions of each type of error were compared for each proficiency group at each of the five error rates. A two-way analysis of variance test was done to determine the degree of inter-group differences in regard to grammatical appropriateness of errors. No significant differences were found between the groups in proportions of non-word responses or of words that violated syntactic context, either between proficiency groups or between error rates. But there were notable differences between the proficiency groups by error rate both for errors that were grammatically inappropriate and for non-word responses. The sum of the proportions over error rates for each type of error is given in Table 16; the raw data, listing numbers of errors of each type at all levels of error rate, is found in Appendix D, Table 25. The analysis of variance is summarized in Table 17.

While differences between groups do not reach statistical significance, there was a tendency for poor readers to make more contextually

TABLE 16

Proportions of Grammatically

Inappropriate and Non-Word Errors

Proficiency Groups	Contextual Errors to Total Substitutions	Non-word Errors to Total Subs.	Total
Good Readers	1.043	2.509	3.552
Average Readers	.925	1.860	2.785
Poor Readers	1.127	1.333	2.460
Total	3.095	5.702	8.797

inappropriate errors and for good readers to make more non-word errors. Whether the greater number of ungrammatical errors among the poor

TABLE 17

Overall Analysis of Variance Based on Proportions

of Grammatically Inappropriate and Non-Word Errors

Source	df	MS	F	P
Context Errors				
Ability	2	1585.866	2.621	NS
Rate	4	1740.566	2.876	$p<.10$
Error	8	605.177		
Total	14			
Non-word Errors				
Ability	2	2257.266	0.879	NS
Rate	4	9099.496	3.545	$p<.10$
Error	8	2566.851		
Total	14			
Total				
Ability	2	29012.561	0.647	NS
Rate	4	99732.690	2.225	$p<.10$
Error	8	44832.808		
Total	14			

readers can be taken as a sign that the poor readers have a greater tendency toward ungrammatical responses is doubtful when one notes that the better readers are prone to give non-word responses. This observation obscures the issue of grammatical appropriateness. While one would probably not defend a non-word as grammatically appropriate, there is no assurance that the child is unaware of grammatical constraints when he makes non-word errors. Some subjects, in fact, created non-words with suffixes that made the responses seem to "fit" the context--for example, the third-person singular -s on non-words in the verb position. The problem of non-words also obscures the relationship between response to the graphic display and response to syntactic context as a function of reading difficulty. An inverse relationship between these two variables was suggested in Chapter I: as it became necessary to give increased attention to processing grephemes, it seems that the reader's sense of grammatical context would decrease, resulting in a larger number of ungrammatical errors at the higher levels of error rate. How many of the non-word errors were made with awareness of grammatical constraints cannot be determined.

The large proportion of non-word responses made by good readers seems to indicate greater reliance on grapheme-phoneme correspondences. When he was confronted with a difficult word, the good reader in this study apparently preferred to create a graphically and phonetically appropriate response, disregarding word meaning and syntactic context. The poor reader, on the other hand, exhibited a greater tendency to respond with a legitimate word. It is possible that these legitimate

words might have been appropriate to the meaning of the word missed or to the passage as a whole--a problem beyond the scope of this study.

Hypothesis 5

> Hypothesis 5: Tendency to read by phrases does not differentiate readers of different levels of reading proficiency.

Before actually testing the hypothesis, it was decided that a study of subjects' tendencies to make errors at the beginning of phrases was necessary. This type of analysis was considered an appropriate procedure for determining awareness of grammatical structure. While Johnson (e.g., 1965, 1968) convincingly demonstrated in a paired-associate experiment that his university subjects tended to read by phrases, the writer was hesitant to generalize these results to third-grade youngsters in an oral reading error study. So a binomial expansion test was done to determine the extent to which children also read by phrases. Errors in all passages at the 0-4.9% error rate were tallied in a dichotomous fashion--errors located at the beginning of phrases and errors located elsewhere in the phrases. The raw data for each subject are contained in Appendix D, Table 26. The tendency for errors to occur at the beginning of phrases was significant ($Z=5.83$, $p<.001$). Thus it was concluded that third-grade children, at least those in the population used in this study, do tend to read by phrases and that study of transitional error probabilities is an appropriate method for assessing children's use of phrase structure.

In testing the hypothesis, a chi-square test was used to find

differences between proficiency groups in tendency to read by phrases, as indicated by the location of errors in phrases in low error rate passages. Tendency for errors to occur at the beginning of the phrase and elsewhere in the phrase are summarized in Table 18. The data

TABLE 18

Location of Errors in Phrases

Location of Errors	Proficiency Groups Good	Average	Poor	Total
Errors at beginning	239	108	68	415
Errors not at beginning	164	53	46	263
Totals	403	161	114	678

failed to reject the null hypothesis ($x^2=3.045$, df=2), substantiating the prediction that reading proficiency is not a factor in awareness of phrase structure and in reading by phrases. The data of only five subjects in the groups of thirty (See Appendix D.)--three of the good readers, one of the average readers, and two of the poor readers--failed to conform to this expectation. It seems likely that the apparent failure of the five non-conforming readers to make use of phrases structure stems from factors unrelated to reading proficiency.

Interphrase associations might account for the instances in which subjects do not seem to be "cuing in" to phrase structure. From his study of the effect of association between phrases, Rosenberg (1968) concluded that, while phrases are apparently the means for grouping in materials with low interphrase association, the reading unit may transcend the phrase in high association passages. To

test this hypothesis, Rosenberg designed a sentence-learning task, using words with high association in the same sentence, e.g., the old king ruled wisely. Sentences proclaiming that kings rule, dogs bark, authors write, and surgeons operate would surely have a high subject-verb association; the action described in the verb is the action associated with the noun. These high semantic associations apparently resulted in a low probability of transitional errors occurring between noun phrase and verb phrase. Schlesinger (1968) also noted that high semantic associations are generally stonger than the effect of syntactic divisions.[6] One might, of course, argue that such high-association utterances do not constitute meaningful communication or that they are not typical of English sentences generally. Yet these findings might account for the tendency of some readers not to read by phrases, particularly in passages in which association happens to be strong.

Error position within the clause in passages at the 0-4.9% error rate was also studied, though no hypotheses were made relevant to it. The likelihood that there would be a tendency toward more errors at the beginning of a clause, fewer in the middle, and yet fewer at the

[6] Johnson (1966) also tested for the effect of interphrase association and, contrary to Rosenberg's findings, observed that the syntactic effect was stronger than the effect of semantic association--even in sentences in which semantic association was strong. This contradiction is probably best explained by the different methods used for establishing word associations. While Rosenberg used "natural" associations that the reader would presumably have formed on his own over the years, Johnson attempted to create such associations through repeated presentations of word pairs.

end was suggested by the results of Weber's first-grade study (1970b). While the observations in this regard in the present study were not clear-cut, there was a tendency in the predicted direction, as the data in Table 19 seem to indicate. For each proficiency group the largest percentage of errors occurred in the first third of the clause while the final third of the clause in each group contained

TABLE 19

Percentage of Errors

by Third of Clause

Proficiency Groups	Third of Clause		
	1	2	3
Good Readers	41.3%	31.9%	26.8%
Average Readers	42.4%	27.6%	30.0%
Poor Readers	41.9%	24.8%	33.3%

the smallest percentage of errors. The tendency was strongest among the good readers. But only the good readers exhibited the progressively decreasing number of errors through all three divisions of the sentence--the behavior one would expect if, in fact, reading does proceed in terms of clauses. The data for all subjects are found in Appendix D, Table 27.

The observations of greater frequency of errors at the beginning of clauses than in other locations of the clauses would seem to indicate that reading sometimes is done in terms of clauses--at least in some readers. The fact that the tendency is stronger among the good

readers would seem to suggest that it is the better reader who transcends the phrase in his reading; it would seem that the better the reader, the larger would be the sequences of words by which he reads. The probability that readers read by units of varying size is in accord with the model of the reading process proposed by Venezky and Calfee (1970) and with the research of Levin and Turner (1966). In both studies it is suggested that the phrase is generally the basic linguistic segment of reading but that length of phrase differs according to reader proficiency and difficulty of the reading material. But the problem of segments by which children read must be tested in more highly controlled studies; clause length and clause structure should be controlled.

Hypothesis 6

> Hypothesis 6: Tendency to repeat word sequences by phrases does not differentiate readers of different levels of reading proficiency.

The assumption underlying Hypothesis 6--that repetition of sequences of words conform to phrase structure grammar--was tested for all subjects as a group by means of a binomial expanison test. All repetitions of two or more words at all levels of difficulty were included except for repetitions that would unfairly bias the results in favor of the hypothesis (See Chapter II, pp. 63-64). The data for this measure is found in Table 20. The test indicated that tendency for repetitions to conform to phrases was significant ($Z=13.3$, $p<.001$), showing that examination of repetitions, like examination of error position in a phrase, is an appropriate means

of assessing subjects' use of phrase structure in reading. The raw data are given in Appendix D, Table 28.

TABLE 20

Conformity of Multi-Word

Repetitions to Phrase Structure Grammar

Types of Repetitions	Proficiency Groups			
	Good	Average	Poor	Total
No. Conforming Repetitions	138	104	69	311
No. Non-conforming Repetitions	46	46	25	117
Total	184	150	94	428

Comparisons between proficiency groups regarding conformity of repetitions to phrase structure were made by means of a chi-square test, for which the data is contained in Table 20. The data supported the null hypothesis (x^2=1.256, df=2) that reader proficiency is not a variable in tendency to repeat words by phrases--further evidence, it would seem, that readers of all levels of proficiency observed in this study were well aware of phrase structure in their reading.

The results of analyses of data pertinent to hypotheses 5 and 6, both of which claim equal awareness of phrase structure across ability groups, seem to be supported by the data in Hypothesis 1 (See pp. 73-75.). There it was noted that the subjects of all three proficiency groups made similar types of errors on short function words; these errors were markedly different from errors made on content words of the same length. This observation was taken as an indication that a all subjects regard structure words differently from content words--

a prerequisite, presumably, for awareness of phrase structure and sentence structure.

SUMMARY

In this chapter the results of the tests of the six hypotheses have been reported and discussed. Hypothesis 1, pertaining to location of error within a word, was accepted with qualification; while the tendency for long words was statistically significant in the hypothesized direction, an opposite tendency was seen in short word errors. Hypothesis 2, which predicted increasing graphic similarity of erroneous responses to their stimuli as reading difficulty increased, was also accepted. But Hypothesis 3, pertaining to graphic similarity between proficiency groups with error rate held constant, was rejected; while there were tendencies in the hypothesized direction, the differences were not significant. The three null hypotheses addressed to subjects' use of syntactic constraints were accepted. There were no significant differences between proficiency groups in tendency for errors to be grammatically appropriate, nor did measures of repetitions to phrase structure grammar show any differences between groups in awareness of, and response to, sentence structure.

Some implications of these results for teaching and for research are suggested in the following chapter.

Chapter IV

SUMMARY, LIMITATIONS, AND IMPLICATIONS

Summary

Problem

The study was designed to find differences in the reading strategies of good, average, and poor readers through an analysis of substitution errors made in oral reading. Specifically, the extent to which good, average, and poor readers depend upon the two main types of reading cues--the graphic features of the word and the context in which the word appears--was examined. Substitution errors were the center of focus since they would seem to yield the greatest amount of insight into reading strategy.

Generally speaking, there are two explanations, not necessarily mutually exclusive, for differences in performance between readers of different levels of reading proficiency. One view is that there is a qualitative, as well as a quantitative, difference in the errors of different types of readers. If this is true, there should be a weakness, or perhaps a syndrome of weaknesses, that characterize the oral reading errors of poor readers--or certain strengths found in the errors of good readers. On the other hand, Malmquist, in examining the errors of readers of different levels of proficiency, found no difference in quality of errors; good, average, and poor readers in his study made about the same proportion of substitutions, omissions,

additions, and reversals. The difference, he concluded, is a quantitative one: the better the reader, the smaller the number of errors made. Malmquist did not, however, examine errors for degrees of appropriateness of erroneous response to the printed word. The writer believed that, by using some linguistic measures of graphic-phonetic and syntactic resemblance of stimulus and response, it might be possible to find qualitative, as well as quantitative, differences between the errors of readers of different levels of reading proficiency.

Subjects

The study was done at third-grade level, with thirty children assigned in equal numbers by sex to three groups according to reading proficiency. Criteria for placement in proficiency groups were reading achievement test performance and teacher evaluation of each child's reading proficiency.

Procedures

The oral reading errors of good, average, and poor readers were examined for graphic-phonetic and syntactic resemblance to the printed words. In order to hold error quantity differences between groups constant, reading materials were chosen at different levels of vocabulary and syntactic difficulty so that all subjects could be compared at the same levels of error rate. These materials were chosen on the basis of publishers' ratings, evaluation by readability formulas, and the results of a pilot study. In a series of fifteen-minute tasks, each subject began reading first-grade level materials and continued through increasingly difficult readings to the point at which he made over twenty errors for every hundred words he read. Thus word

difficulty could be examined across and within proficiency groups with the effect of error rate held constant.

Two measures were used to determine subjects' dependence on grapheme-phoneme relationships. An examination of stimulus and response was made for location of error within the word. Every substitution error was graded according to the location of the error in the word in which it occurred: initial phoneme, middle phoneme(s), and final phoneme. Words on which substitutions were made were classified by length of word. Words of five or more phonemes were considered long words. Short words were those words that contained at least three letters and three or four phonemes. Words having less than three phonemes were omitted from this part of the analysis. Short words were subgrouped into function words and content words in addition to being examined for match in the first two letters and the last two letters. The Graphic Similarity Index developed by Weber takes into account the number of letters that match in stimulus and response as well as similarity of word length. Values are assigned by the formula to the various positions within the word where there is a match. These numerical values are intended to parallel the relative importance of the different positions within the word as cues for word recognition.

Two measures were used to assess subjects' awareness of syntax. The first of these was an evaluation of the appropriateness of substitution errors to the sentences in which they occurred. If the substitution fit the sentence grammatically, it was inferred that the subject was sensitive to sentence structure. An error was considered gram-

matically appropriate if it "fit" the sentence from beginning to point of error. Another means of determining use of sentence structure was an analysis of the position of children's errors in phrases when they were reading at a low error rate. Psycholinguistic research has shown that phrases (as determined by immediate constituent analysis) are the units by which readers decode. If this is so, one would expect more errors to occur at the beginning of phrases. Grammatical context is least restrictive at that point; hence there should be more errors at the beginning of a phrase. Furthermore, repeated sequences of words were evaluated for conformity to phrase structure grammar. Since the purposes of a child's repeating more than one word apparently is an attempt to strengthen his sense of the context in order to deal with a troublesome word or sequence of words, it would seem that he should, when he repeats a sequence, regress to the beginning of a phrase. Only those repeated sequences in which a child had the option of choosing a word that was not located on a phrase boundary were tallied.

Results

Analysis of errors on the word level generally showed that the better the reader, the greater was the graphic-phonetic conformity of his mismatches to the printed word. The measure of error location in a word indicated that on longer words the better readers were more likely to respond with a word that matched the stimulus in both initial and final positions. In rank ordering the error types on long words for each subject it was found that the Type 4 error--the error type that showed greatest attention to the graphic display--was the most

common error for all subjects in the high group while seven of the average and five of the poor readers made this error most frequently. But on short content words the poorer readers made the largest proportion of graphically appropriate errors (Type 4 Error) while the better readers demonstrated less attention to the graphic display in their errors. Subjects in all three groups showed a tendency to deal with short function words in a markedly different manner than with short content words or long words. Proportions of error types did not differ significantly according to error rate. The graphic similarity measure developed by Weber revealed that difficulty of the material affected performance similarly in all three groups. As the difficulty of the material increased, so did the graphic similarity scores. This measure also showed that the better the reader, the higher his graphic similarity score tended to be--at any particular level of error rate--though this difference did not reach statistical significance.

The results of the syntactic measures reveal that the readers of all three proficiency groups were highly sensitive to sentence structure. The errors of all three ability groups showed a strong tendency to be grammatically appropriate. The differences between the groups were not significant, though the better the reader, the greater was the tendency for errors to be syntactically appropriate. However, the better readers showed a greater tendency to respond with non-words. As error rate increased, the percentage of grammatically appropriate substitutions dropped in all three groups. There was no difference between the groups in tendency to read by

phrases according to both measures of phrase reading--location of errors at the beginning of phrases and conformity of repetitions to phrase structure grammar. In all three proficiency groups errors in passages in which the subject achieved a low error rate tended to be located in phrase boundaries. Also, multi-word repetitions in all three groups generally corresponded to linguistic phrases.

This study was an attempt to deal with the question Malmquist (1958) addressed himself to over a decade ago--the differences in the oral reading errors of children of different reading proficiency levels. The conclusions made herein will be found to differ somewhat from those of previous analyses of oral reading errors beyond the first grade. This difference can be attributed largely to use of recently developed linguistic analyses, some of which had not been applied in this type of research. The writer has demonstrated that there are some differences in appropriateness of errors to their stimuli, specifically on the word level. Syntactically, there seems to be little difference in the appropriateness of errors of the three groups. Malmquist concluded that the errors of readers of different levels of reading proficiency differ only in degree, not in kind; poor readers make the same kinds of errors as good readers, but in larger quantity. Malmquist, however, was primarily interested in proportions of error types--substitutions, omissions, additions, and the like--rather than in appropriateness of errors to the stimulus words. While the observations of the present study agree essentially with Malmquist's observations on the syntactic level, the data on the word level in this study suggests some intergroup differences.

Specifically, good readers showed more accuracy in responding to the graphic and phonetic features of long words, attending more carefully to the final positions of words.

Limitations of the Study

Various limitations in this study can be cited. Some of these limitations are inherent in oral reading error analyses generally; others are peculiar to this study.

The oral reading error analysis offers numerous advantages and also some limitations. The fact that meaningful reading materials, typical of those in general classroom use, are used as the stimulus in this type of study suggests that there should be a greater transferability of findings and conclusions to the teaching of reading and to models of the reading process than might be true of verbal learning studies using non-words or only a limited range of sentence structures. But in using such "natural" reading materials, the experimenter has less control over his materials--and consequently less control over his subjects' responses to those materials--than in studies in which words or sequences of words are carefully chosen or constructed with specific purposes in mind. It has been pointed out how word structure and word frequency are only grossly controlled in the present study. On the syntactic level, there is no control over length and structure of clauses and phrases, nor over the types of words in these structures. While Goodman's point about the logic of oral reading errors is well taken, there are yet some errors which are difficult to explain. Thus oral reading error analyses should be regarded as rather gross studies. It might be that oral reading studies should

be the starting point and the culmination of many linguistic aspects of reading research. That is, initially the oral reading error analysis might be used as a "feeler," to point out areas in which more controlled research should be done. At the other end of the continuum this type of inquiry should provide some insight into how valid the findings of the more closely controlled studies are in more general reading situations.

The problem of the relationship of oral reading performance to the silent reading process requires that the results of all oral error analyses be taken with some qualification. While Gilmore (1947) presents convincing evidence that the correlation between the two modes of reading is a high one, it is not justifiable to take one as a carbon copy of the other. One can claim only that the oral reading error analysis is probably one of the best means yet devised for studying the silent reading process in a natural reading situation.

Given the limitations of the oral reading analysis, there are certain limitations peculiar to the present study. While the corpus of errors analyzed is indeed substantial, the number of subjects making those errors is relatively small--a problem in in-depth error analyses generally. To some extent, at least, size of population is sacrificed for depth of analysis. Also, the fact that all subjects were drawn from one school system means that all subjects were taught reading in essentially the same teaching method. The writer expects that there would be some differences in performances of children taught by different methods.

Finally, the subject's familiarity, or lack of it, with the subject matter may have influenced responses. While there was an effort to control for cognitive difficulty, no guarantee can be made that such control was actually achieved. It is conceivable that students' familiarity with the topic of a selection could have some effect on decoding behavior.

The implications suggested below should be considered with these limitations in mind.

Implications for Teaching

Many of the results of the study, particularly on the word level, do not seem directly applicable to the teaching of reading at this point. Further research is needed to determine specific areas of strength and weakness in decoding. The measures on the word level tended to distinguish the good reader from the average and poor reader. Most of the measures of location of error in the word reveal what characterizes the good reader. Further research into such errors might well uncover specific graphemes or combinations of graphemes that prove troublesome for the poor reader -- finding that could have significant implications for remedial teaching. The measures at this level give some suggestion that the poor reader's problem is largely one of decoding. But sight vocabulary also appears to be an area of weakness, as suggested by the great care with which poor readers attack short content words, many of which are high in frequency. Yet the poor readers' errors on short function words would seem to indicate that they do have at least the most frequent words of the language in their sight vocabularies.

On the sentence level, it was observed that the syntactic appropriateness of substitution errors decreased with increase of error rate, a trend that was more noticeable among the poor readers. This observation seems to provide some experimental evidence for the insistence by reading educators that poor readers--all readers, for that matter--should not be assigned selections that are beyond their reading ability. The poor reader's problem in this study seemed to be one of decoding; the results show that he appears to be well aware of grammatical appropriateness and sentence structure--when he reads at a relatively low error rate. But when reading selections are unduly frustrating, children are more likely to lose "sentence sense"--in addition to having decoding difficulties. The Graphic Similarity Index and the measure of grammatical appropriateness of errors revealed that the hypothesized relationship between attention to the graphic display and awareness of grammatical context is valid, as measured by grammatical appropriateness of errors. As attention to the graphic display increases, awareness of grammatical context decreases. To read without the necessary decoding skills is a real problem in itself; to read, in addition, without being certain of syntactic context must be very frustrating indeed. It would seem, too, that as the child's sense of syntax weakens, so would his awareness of meaning. Thus it seems important that teachers individualize reading assignments, keeping the ability of each child in mind; expecting children to read materials that are beyond their decoding ability is educationally unsound.

The results of this study seem to have implications regarding the teaching of sentence structure for remediation of reading difficulties, as has been suggested in some literature. Lefevre (1965) in particular has directed attention to the teaching of sentence structure generally:

> Greater attention should be given to developing sentence sense in reading and writing and less to learning individual words. It is probable that [sic] given a mastery of basic sentence structure, vocabulary would take care of itself.... (p.23)

Appropros to this suggestion, Lefevre characterizes the poor reader not as one who lacks an adequate sight vocabulary or has failed to master the grapheme-phoneme relationships. Rather, the poor reader lacks "sentence sense":

> Lacking a sure grasp of the printed sentence as the common building block of written discourse, the crippled reader cannot comprehend what he 'reads' as organized, coherent form. Instead, he tends to register only arbitrary, random elements, and even to miss important language structures altogether in the material the writer sets before him. He sees a subject without its verb, a verb without its subject; he combines subjects with the wrong verbs and verbs with the wrong subjects; he attaches expanding phrases to the wrong sentence elements, or 'reads' them by themselves, without any structural context. (p. 23)

Thus it is apparent that Lefevre takes issue with those reading specialists who regard phonics instruction as a basic component of a beginning reading program and as one of the most useful means of dealing with reading disability.

This study offers practically no support for Lefevre's position on reading disability. When readers of the three proficiency groups

are compared on the measures of syntactic awareness, there appears to be no difference in the syntactic appropriateness of their errors.[5] The tendency for readers to make errors at the beginning of phrases, where syntactic context is least restrictive, and their tendencies to make multi-word repetitions corresponding to phrases were nearly equal across proficiency groups. Also, an overwhelming number of substitution errors conformed to the sentence structure. "Sentence sense," then, was not a discriminating factor in the readers of this study, a finding that is in essential agreement with error analysis studies of first-graders in which reading ability has been a variable (Clay, 1968; Weber, 1970b). Recognition of the grammatical meaning of these words would seem a prerequisite for perceiving grammatical relationships--as Lefevre rightly insists. The similar behaviors of all three groups of subjects on short function words, as compared to their performance on short content words, would seem an indication that all subjects regarded these structures in much the same way--as words of grammatical importance. Lefevre's emphasis on the importance of the reader's knowledge of the structure of the language is surely valid. However, as this writer has stressed in Chapter I, it should not be necessary to <u>teach</u> the structure of the language to children who have been using language for five years or more; they are already familiar with its basic structure. Hence it is doubtful that teaching devices designed to encourage reading by phrases are necessary in the teaching

[5] While the poor readers made more syntactically inappropriate errors on real words, the large number of non-word errors made by the good readers obscures the problem of syntactic appropriateness of errors.

of reading to children who speak standard English. In short, it seems that if there is one problem the poor reader does not have, it is awareness and response to the structure of English. This is not to say that this problem is not a concern in teaching reading to students who are linguistically different from the population of this study--for example, children who speak a non-standard dialect or are learning English as a second language.

Finally, the results of this study seem to indicate that a reasonably strong phonics program does not necessarily produce word-by-word readers. This is not to say that an overemphasis on phonics skills might not be harmful. These subjects were taught in a basal reader program with a strong phonics supplement. While they were very conscious of the parts of words, the syntactic appropriateness of most of their errors and the tendency to make use of phrase structure leave little doubt that they were also very much aware of sentence structure.

Implications for Future Research

Word Level

A replication of the study with variation in subject population could yield valuable results. For example, the basic design and analysis could be applied to the subject of dialect differences and the problems these differences cause in the teaching of reading. A study of the oral reading errors of urban black subjects, for example, might serve to confirm or disprove some of the claims made by linguists about phonetic, morphological, and syntactic differences between black English and standard English. This type of study could also be used

to examine the problem of methods of teaching reading to children who speak a non-standard dialect. A replication of this study using subjects who have been taught reading in a predominantly word method program might well reveal differences in reading strategy between children who have been taught by both methods. Do children who have been taught in a systematic phonics program depend on grapheme-phoneme relationships to a greater extent than do children taught in a predominantly word-method program? Do word-method subjects demonstrate a greater tendency to rely on their sight vocabularies than on phonic skills? Are word-method subjects more prone to "cue in" to meaning and give more semantically appropriate responses? A comparison of two such groups in graphic similarity of errors to their stimulus words and in location of error within the word would be likely means of revealing differences, if any exist.

The fact that the errors of poor readers in the study showed a strong tendency to attack even short words carefully suggests another problem regarding method of teaching reading. Poor readers might benefit from a teaching method that provides for more learning of words as wholes. Perhaps a part of the poor readers' problem in this study was an inability "to see the forest because of the trees"; they may have been concerned with word analysis at the expense of seeing words as whole units. An examination of oral reading errors might reveal the extent to which poor readers taught in other reading methods have more limited sight vocabularies. This might be a problem with poor readers who have been taught reading in a systematic phonics program. On the other hand, it might be a problem that poor readers generally face.

Further research into oral reading errors might also reveal particular letters or patterns of letters that prove especially troublesome for poor readers. Perhaps they have trouble in distinguishing vowels of alternate pronunciations in different environments or in processing consonant or vowel clusters. The observation that in the present study vowel sounds were faulty in considerably more responses made by poor readers than of good readers seems worthy of further attention. The identification of such patterns would surely have implications for remediation procedures. On the other hand, it might well be found that word difficulties are strictly an individual matter, with different readers having problems with different sounds or sound patterns.

For further investigation of sight vocabulary in children it seems prudent to refine the method of specifying word frequency in this study--the function word-content word distinction. Examining differences in word attack strategy according to word frequency as specified by the Kučera and Francis frequency list (1967) should give a more accurate picture of the relationship between word attack and word frequency. Such a study would be done under the assumption that a strong relationship exists between word frequency and sight vocabulary. Such a study might also indicate why good and average readers attacked short content words and long words so differently in the present study--whether these differences are a function of word length or word frequency, or both.

A longitudinal design of the present study or repetition of it at various age levels would be a means of further examining an observa-

tion made by Venezky, Chapman, and Calfee (1970)--that the location of errors in children's pronunciation of synthetic words varied with age. The older the children, the greater was their tendency to make more graphically accurate errors--errors in which printed word and erroneous response matched in both initial and final positions. An oral reading error analysis carried out either on poor readers of different ages or on the same poor readers over a period of years should indicate the extent to which the results of Venezky, Chapman, and Calfee hold true in a sentence context. Perhaps the poor reader's perception of, and response to, words improves with age. The results of such a study might lend support to the proposal advanced recently by some reading specialists that training in grapheme-phoneme relationships be delayed for some immature children.

Syntactic Level

It seems advisable to make more controlled studies of readers of differing levels of ability in their response to syntactic structures. Sentence-learning tasks similar in design to those done by Johnson might be used to further test hypotheses regarding syntactic cues in this study--for example, the extent to which length of phrase and length of clause affect use of linguistic segments by different types of readers and the effect of noun or article in the initial position of clause or phrase. Gradating stimulus materials for vocabulary and syntactic difficulty would be done as it has been done in the present study; results can be considered valid indicators of use of syntactic structures only when readers are operating at a relatively low error rate.

The results of the sentence level aspect of the study show a tendency for readers of all three levels of ability to read by phrases. This has previously been demonstrated with adult readers in more controlled studies (e.g., Johnson, 1965, 1966, 1968; Levin and Turner, 1966; Schlesinger, 1968). The fact that intergroup differences were minimal was taken as evidence that children of various levels of reading proficiency were aware of constituent structure. But this tendency was not found in all readers in this study. It was suggested in the previous chapter that association between words separated by phrase boundaries may have lessened the tendency to read by phrases. A sentence-learning task similar to Rosenberg's (1968), with distinctions made between high association and low association word relationships, might reveal the extent to which word association influences reading in young readers generally--and among good, average, and poor readers specifically.

The measure of reading by clauses indicated that some readers may indeed tend to read by these larger structures. The tendency appeared to be stronger among the good readers. A study in which clause length and structure are carefully controlled could yield information pertinent to this question.

Finally, it seems that some of the syntactic measures developed in the course of this study should be useful tools for future reading research. Measures of location of error in linguistic units, adapted from the work by Johnson and Weber, might prove to be a valuable means of assessing response to sentence structure in future research. Conformity of repetitions to phrase structure grammar proved to be

another effective indicator of the use of syntactic structures in reading.

It is hoped that the results of this study will encourage further research in linguistic aspects of reading behavior and in the problems of individual differences in reading.

BIBLIOGRAPHY

Allen, P. A psycholinguistic analysis of the substitution of miscues of selected oral readers in grades 2, 4, and 6, and the relationships of these miscues to the reading process, a descriptive study. Unpublished doctoral dissertation, Wayne State University, 1969.

Bennett, A. An analysis of errors in word recognition made by retarded readers. Journal of Educational Psychology, 1942, 33, 25-38.

Bever, T. G., Lackner, J. R., & Kirk, J. R. The underlying structures of sentences are the primary units of immediate speech processing. Perception and Psychophysics, 1969, 5, 225-234.

Biemiller, A. The development of the use of graphic and contextual information as children learn to read. Reading Research Quarterly, 1970, 6, 75-96.

Bloomfield, L. Language. New York: Henry Holt, 1933.

Brown, R. & Bellugi, U. Three processes in the child's acquisition of syntax. Harvard Education Review, 1964, 34, 133-151.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: The MIT Press, 1965.

_____. Syntactic Structures. The Hague: Mouton and Co., 1957.

Clay, M. M. A syntactic analysis of reading errors. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 434-438.

Dale, E. & Chall, J. S. A formula for predicting readability. Educational Research Bulletin, Ohio State University, 1948, 28, 11-20, 28, 37-54.

Daw, S. E. The persistence of errors in oral reading in grades four and five. Journal of Educational Research, 1938, 32, 81-90.

Duffy, G. B. & Durrell, D. D. Third grade difficulties in oral reading. Education, 1935, 35, 32-40.

Fodor, J. & Bever, T. The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-420.

Francis, W. N. The structure of American English. New York: Ronald Press, 1958.

Fries, C. C. Linguistics and reading. New York: Holt, Rinehart, and Winston, 1962.

_____. The structure of English. New York: Harcourt, Brace and World, 1952.

Garrett, M., Bever, T., & Fodor, J. The active use of grammar in speech perception. Perception and Psychophysics, 1966, 1, 30-32.

Gates, A. I. MacGinitie, W. H. Gates-MacGinitie reading tests. New York: Teacher's College Press, 1965.

Gleason, H. A., Jr. An introduction to descriptive linguistics. New York: Holt, Rinehart and Winston, 1961.

_____. Linguistics and English grammar. New York: Holt, Rinehart, and Winston, 1965.

Gilmore, J. V. The relationship between certain oral reading habits and silent reading comprehension. Unpublished doctoral dissertation, Harvard University, 1947.

Goodman, K. S. Analysis of oral reading miscues: applied psycholinguistics. Reading Research Quarterly, 1969, 5, 9-30.

_____. The psycholinguistic nature of the reading process. In K. S. Goodman (Ed.), The Psycholinguistic Nature of Reading Process. Detroit: Wayne State University Press, 1968. Pp. 43-57.

_____. A psycholinguistic view of reading comprehension. In New Frontiers in College-Adult Reading. Fifteenth Yearbook of the National Reading Conference. Milwaukee: The National Reading Conference, 1966.

_____. Reading: a psycholinguistic guessing game. Journal of the Reading Specialist, 1967, 6, 126-135.

Goodman, Y. M. A psycholinguistic description of oral reading phenomena in selected young beginning readers. Unpublished doctoral dissertation, Wayne State University, 1967.

Hardyck, C. D. & Petrinovitch, L. F. Treatment of subvocal speech during reading. Journal of Reading, 1969, 12, 361-368, 419-422.

Ilg, F. L. and Ames, L. B. Developmental trends in reading behavior. Journal of Genetic Psychology, 1950, 76, 291-312.

Johnson, N. F. The influence of associations between elements of structured verbal responses. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 369-374.

_____. The influence of grammatical units on learning. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 236-240.

_____. The psychological reality of phrase-structure rules. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 469-475.

Kenyon, J. S. & Knott, T. A. A pronouncing dictionary of American English. Springfield, Mass.: G. and C. Merriam, 1953.

Kučera, H. & Francis, W. N. Computational analysis of present-day American English. Providence, R. I.: Brown University Press, 1967.

Lees, R. B. The grammar of English nominalizations. Bloomington: Indiana University Press, 1967.

Lefevre, C. A. Linguistics and the teaching of reading. New York: McGraw-Hill, 1964.

Levin, H. and Turner, E. A. Sentence structure and the eye-voice span. Project Literacy Reports, No. 7. Ithaca: Cornell University, 1966.

Madden, M. and Pratt, M. An oral reading survey as a teaching aid. Elementary English Review, 1941, 18, 122-126, 159.

Malmquist, E. Factors related to reading disabilities in the first grade of elementary school. Stockholm: Almqvist & Wiksell, 1958.

Marchbanks, G. & Levin, H. Cues by which children recognize words. Journal of Educational Psychology, 1965, 56, 57-61.

Page, W. D. A psycholinguistic description of patterns of miscues generated by a proficient reader in second grade, an average reader in fourth grade and an average reader in sixth grade encountering ten basal reader selections ranging from pre-primer to sixth grade. Unpublished doctoral dissertation, Wayne State University, 1970.

Payne, C. S. The classification of errors in oral reading. Elementary School Teacher, 1930, 31, 142-146.

Pike, K. L. Taxemes and immediate constituents. Language, 1943, 19, 65-82.

Rosenberg, S. Association and phrase structure in sentence recall. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 1077-1081.

Schale, F. C. Changes in oral reading errors at elementary and secondary levels. Unpublished doctoral dissertation. University of Chicago, 1966.

Schlesinger, I. M. Sentence structure and the reading process. The Hague: Mouton, 1968.

Schummers, J. L. Word pronunciation in the oral sight-reading of third-grade children. Unpublished doctoral dissertation, University of Minnesota, 1956.

Shapiro, S. I. & Gordon, G. P. Contemporary norms of word and phonetic frequencies. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 92-94.

Sokal, R. R. & Rohlf, F. J. Biometry: the principles and practice of statistics in biological research. San Francisco: W. H. Freeman, 1969.

Spache, G. D. A new readability formula for primary-grade reading materials. Elementary School Journal, 1953, 53, 410-413.

Swanson, D. E. Common elements in silent and oral reading. Psychological Monographs, 1937, 48 (215), 36-60.

Venezky, R. L., Chapman, R., & Calfee, R. C. The development of letter-sound generalizations, second grade through fourth grade. Working Paper from the Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, Wisconsin, in press.

Venezky, R. L. & Calfee, R. C. Models for the reading process. Working Paper from the Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, Wisconsin, June, 1970, No. 38.

Weber, R. First-graders' use of grammatical context in reading. In H. Levin & J. P. Williams (Eds.), Basic Studies in Reading. New York: Basic Books, 1970(a)

_____. A linguistic analysis of first-grade reading errors. Reading Research Quarterly, 1970(b), 5, 427-451.

Appendices A through D have been omitted from this publication,
but are available on microfilm from Memorial Library,
University of Wisconsin, Madison, Wisconsin.